An overview of methods for estimating cancer patient survival

Paul W Dickman

King's College London
29 June 2021
http://pauldickman.com/talk/

Karolinska Institutet

Overview of today's talk

Please interrupt!

- About me and my department.
- Measures used in cancer control; why study patient survival?
- Intro to net/relative survival and why it is the measure of choice for estimating patient survival using registry data.
- 'Real-world' alternatives to net survival; crude survival.
- Loss in expectation of life (time permitting).
- Slides at: `https://pauldickman.com/talk/cancer-survival-methods-kcl-2021/`

About me

- Born in Sydney Australia;
  studied mathematics and statistics in Newcastle (Australia).
- Worked in health services research;
  dabbled in industrial process control and quality improvement.
- Arrived in Sweden November 1993 for a 10 month visit to cancer epidemiology unit at KI. Stayed in Sweden for most of my PhD.
- Short Postdoc periods at Finnish Cancer Registry
  and Karolinska Institutet (cancer epidemiology).
- Joined current department in March 1999, attracted by the strong research environment and possibilities in register-based epidemiology.

My research interests

- Development and application of methods for population-based cancer survival analysis, particularly the estimation and modeling of net survival.
- General interest in statistical aspects of the design, analysis, and reporting of epidemiological studies.
- Epidemiology, with particular focus on cancer epidemiology and perinatal/reproductive epidemiology.
- Lots of administrative work.

I found paradise! [1]

A paradise for epidemiologists?

Hans-Olov Adami The Lancet 1996;2:588

For three reasons—the structure of its health system, the existence of nationwide registers, and the systematic use of national registration numbers—Sweden offers exceptional opportunities for epidemiological research.

- I would add 'willingness of the public to contribute to research' and 'outstanding clinical researchers'.

About Karolinska Institutet and MEB

- Karolinska Institutet is a medical university in Stockholm;
  - 6500 undergraduate students;
  - 2000 doctoral students;
  - 5000 FTE staff;
  - 45th overall (1st in Sweden) in Shanghai rankings.
- Department of Medical Epidemiology and Biostatistics (MEB) one of 22 departments; 300 staff (including doctoral students).
- Established in 1997 when Department of Cancer Epidemiology moved from Uppsala University.
- Focus on register-based epidemiology; especially strong in cancer epidemiology, psychiatric epidemiology, and biostatistics.

My focus is population-based cancer patient survival

- The term ‘population-based’ refers to the fact that we are estimating survival for all patients in a geographically-defined population (i.e., from a population-based cancer registry) rather than, for example, patients enrolled in a clinical trial.
- Population-based studies of patient survival provide a measure of the effectiveness of the health care system in diagnosing and treating those cancers that arise in the entire population.
- Note that this includes the actions of the health care system in promoting public awareness of cancer and the importance of recognising symptoms and consulting a health professional when symptoms occur.

Population-based measures used in cancer control

- The key measures are incidence, mortality, and survival.
- By 'mortality' we typically mean mortality in the population, whereas 'survival' is nothing more than mortality among those diagnosed with cancer (transformed to the survival scale).

$$S(t) \;=\; \exp\left(-\int_0^t h(u)\,\mathrm{d}u\right) = \exp\left(-H(t)\right)$$

where $H(t)$ is the *integrated hazard* or *cumulative hazard*.

- We should not study any one of these three measures in isolation; in particular we should consider incidence trends when interpreting trends in patient survival [2, 3, 4, 5, 6].

There is heavy interest in international comparisons For example, CONCORD-3 Study (2018) [7]

Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries

*Claudia Allemani, Tomohiro Matsuda, Veronica Di Carlo, Rhea Harewood, Melissa Matz, Maja Nikšić, Audrey Bonaventure, Mikhail Valkov, Christopher J Johnson, Jacques Estève, Olufemi J Ogunbiyi, Gulnar Azevedo e Silva, Wan-Qing Chen, Sultan Eser, Gerda Engholm, Charles A Stiller, Alain Monnereau, Ryan R Woods, Otto Visser, Gek Hsiang Lim, Joanne Aitken, Hannah K Weir, Michel P Coleman, CONCORD Working Group**

Summary

Background **In 2015, the second cycle of the CONCORD programme established global surveillance of cancer survival as a metric of the effectiveness of health systems and to inform global policy on cancer control. CONCORD-3 updates the worldwide surveillance of cancer survival to 2014.**

Beral & Peto, *BMJ* 2010;341:c4112

UK cancer survival statistics

Are misleading and make survival look worse than it is

RESEARCH, p 335

Valerie Beral professor of epidemiology, Cancer Epidemiology Unit, University of Oxford, Oxford OX3 7LF
pa.valerie.beral@ceu.ox.ac.uk
Richard Peto professor of medical statistics and epidemiology, Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), University of Oxford, Oxford OX3 7LF

Competing interests: Both authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from either

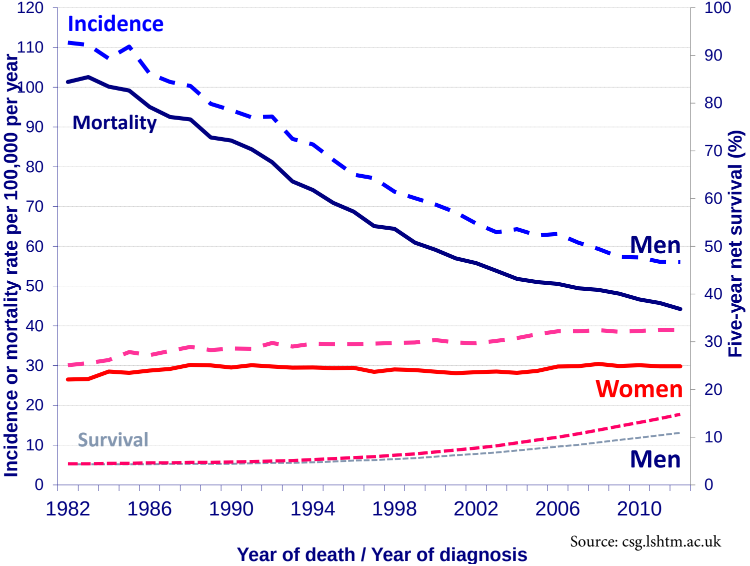
In the linked article, Autier and colleagues report that (population based) breast cancer mortality rates have fallen over the past two decades in many European countries, with a greater decline in the United Kingdom than in any other large country.[1] That the UK is leading Europe in the speed with which national breast cancer mortality rates are falling is in stark contrast to, and at first sight difficult to reconcile with, claims that survival after breast cancer onset is worse in the UK than elsewhere in western Europe.[2]

The unpromising UK cancer survival estimates are, however, misleading. In contrast, population based mortality trends are reasonably reliable (at least in middle age, for example, people aged 35-69 years) because a death certificate is legally required before someone can be buried vival calculations based on registry data make UK cancer survival rates seem significantly worse than they really are.

Information in cancer registries on deaths from cancer is virtually complete because every death certificate that mentions cancer is automatically sent to one of the regional registries that, between them, cover the UK. That cancer is then registered, and further information is sought (not always successfully) from medical records. Death certificates have for decades played an important role in the way UK registries identify people with cancer. Without this source of information, many such cancers could have been missed; even with it, many people who die of cancer may have no record other than the death certificate ever traced by the registry ("death certificate only" cases) or may have had only the later phase

- 'In the absence of internationally comparable data on breast cancer survival rates, it is of interest to compare the reliably known trends in population based mortality rates in middle age.'

Lung cancer incidence, mortality and survival (age-standardised) England, 1982-2012, by sex



Source: csg.lshtm.ac.uk

International Cancer Benchmarking Partnership [8, 9]

Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995–2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data

*M P Coleman, D Forman, H Bryant, J Butler, B Rachet, C Maringe, U Nur, E Tracey, M Coory, J Hatcher, C E McGahan, D Turner, L Marrett, M L Gjerstorff, T B Johannesen, J Adolfsson, M Lambe, G Lawrence, D Meechan, E J Morris, R Middleton, J Steward, M A Richards, and the ICBP Module 1 Working Group**

BJC British Journal of Cancer www.nature.com/bjc

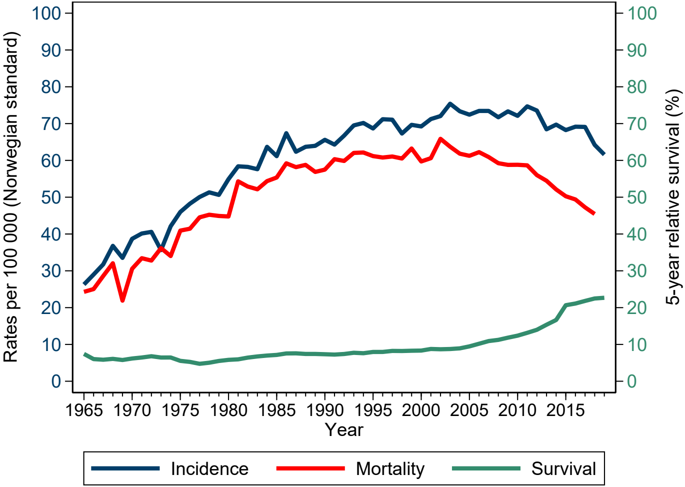
Check for updates

ARTICLE
Epidemiology

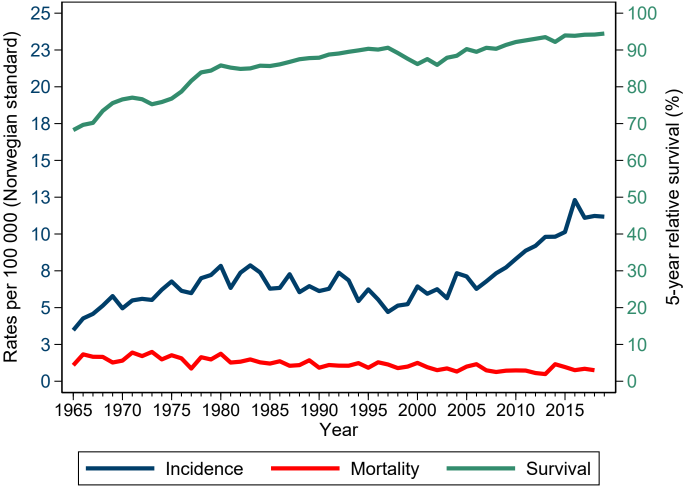
Exploring the impact of cancer registry completeness on international cancer survival differences: a simulation study

Therese M.-L. Andersson[1], Mark J. Rutherford[2,3], Tor Åge Myklebust[4,5], Bjørn Møller[4], Isabelle Soerjomataram[3], Melina Arnold[3], Freddie Bray[3], D. Max Parkin[3,6], Peter Sasieni[7], Oliver Bucher[8], Prithwish De[9], Gerda Engholm[10], Anna Gavin[11], Alana Little[12], Geoff Porter[13], Agnihotram V. Ramanakumar[14], Nathalie Saint-Jacques[15], Paul M. Walsh[16], Ryan R. Woods[17] and Paul C. Lambert[1,2]

Lung cancer, males, Norway [10]



Thyroid cancer, females, Norway [10]



We have a choice of two measures of the probability of death due to cancer

Net probability of death due to cancer = Probability of death in a hypothetical world where the cancer under study is the only possible cause of death

Crude probability of death due to cancer = Probability of death in the real world where you may die of other causes before the cancer kills you

- Net probability also known as the marginal probability.
- Crude probability also known as cumulative incidence function.

How might we measure the prognosis of cancer patients?

- For the moment, I will work on the mortality (among the patients) scale and introduce the two frameworks.
- We could estimate all-cause mortality (among the patients).
- Our interest, however, is typically in mortality associated with a diagnosis of cancer so we often prefer cause-specific mortality.
- When estimating cause-specific mortality only those deaths which can be attributed to the cancer in question are considered to be events.

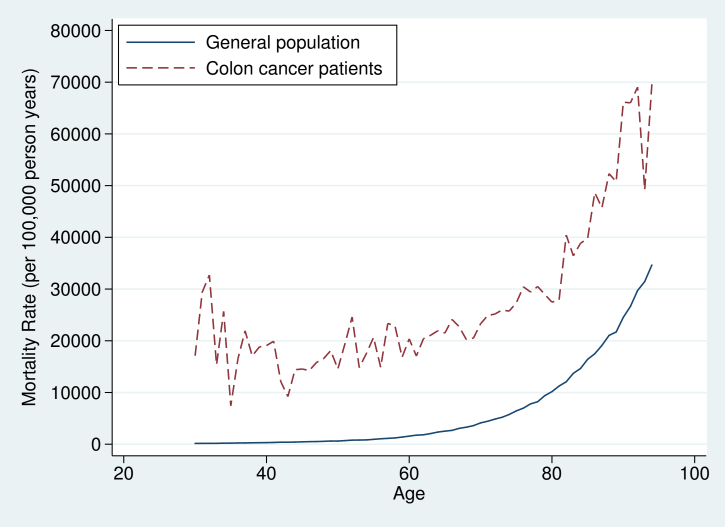
$$\text{cause-specific mortality} = \frac{\text{number of deaths due to cancer}}{\text{person-time at risk}}$$

The survival times of patients who die of causes other than cancer are censored.

Cause-specific survival can estimate net survival (assuming conditional independence)

- Using cause-specific methods requires that reliably coded information on cause of death is available.
- Even when cause of death information is available to the cancer registry via death certificates, it is often vague and difficult to determine whether or not cancer is the primary cause of death.
- How do we classify, for example, deaths due to treatment complications?
- Consider a patient treated with radiation therapy and chemotherapy who dies of cardiovascular disease. Do we classify this death as 'due entirely to cancer' or 'due entirely to other causes'?

All-cause mortality for males with colon cancer and Finnish population



Relative survival aims to estimate net survival (still need conditional independence)

- We estimate excess mortality: the difference between observed (all-cause) and expected mortality.

$$\begin{array}{ccccc} \text{excess} & = & \text{observed} & - & \text{expected} \\ \text{mortality} & & \text{mortality} & & \text{mortality} \end{array}$$

- Relative survival is the survival analog of excess mortality — the relative survival ratio is defined as the observed survival in the patient group divided by the expected survival of a comparable group from the general population.

$$\text{relative survival ratio} = \frac{\text{observed survival proportion}}{\text{expected survival proportion}}$$

Relative survival example (skin melanoma)

Table 1: Number of cases (N) and 5-year observed (p), expected (p^*), and relative (r) survival for males diagnosed with localised skin melanoma in Finland during 1985–1994.

Age	N	p	p^*	r
15–29	67	0.947	0.993	0.954
30–44	273	0.856	0.982	0.872
45–59	503	0.824	0.943	0.874
60–74	449	0.679	0.815	0.833
75+	200	0.396	0.505	0.784

- Relative survival controls for the fact that expected mortality depends on demographic characteristics (age, sex, etc.).
- In addition, relative survival may, and usually does, depend on such factors.

Examples of Relative Survival Being Problematic (Extract from Table 4 from Howlader *et al.* [11])

Selected cancer cohort	White		
	RS, % (95% CI)	CSS, % (95% CI)	Dif., %
Breast			
In situ and <65 y	100.9†	99.7 (99.6 to 99.8)	1.2
In situ and ≥65 y	107.5†	98.6 (98.4 to 98.8)	8.9
Prostate			
Localized/ regional and <65 y	101.3†	98.3 (98.2 to 98.4)	3.0
Localized/ regional and ≥65 y	104.5†	94.8 (94.6 to 94.9)	9.8

Relative survival not as problematic as one might think for lung cancer [12]

Should relative survival be used with lung cancer data?

SR Hinchliffe*,[1], MJ Rutherford[1], MJ Crowther[1], CP Nelson[1,2] and PC Lambert[1,3]

[1]Department of Health Sciences, Centre for Biostatistics and Genetic Epidemiology, 2nd Floor Adrian Building, University Road, University of Leicester, Leicester LE1 7RH, UK; [2]Department of Cardiovascular Sciences, Clinical Sciences Wing, Glenfield General Hospital, University of Leicester, Leicester LE3 9QP, UK; [3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, PO Box 281, Stockholm SE-171 77, Sweden

BACKGROUND: Under certain assumptions, relative survival is a measure of net survival based on estimating the excess mortality in a study population when compared with the general population. Background mortality estimates are usually taken from national life tables that are broken down by age, sex and calendar year. A fundamental assumption of relative survival methods is that if a patient did not have the disease of interest then their probability of survival would be comparable to that of the general population. It is argued, as most lung cancer patients are smokers and therefore carry a higher risk of smoking-related mortalities, that they are not comparable to a population where the majority are likely to be non-smokers.

METHODS: We use data from the Finnish Cancer Registry to assess the impact that the non-comparability assumption has on the estimates of relative survival through the use of a sensitivity analysis.

RESULTS: Under realistic estimates of increased all-cause mortality for smokers compared with non-smokers, the bias in the estimates of relative survival caused by the non-comparability assumption is negligible.

CONCLUSION: Although the assumption of comparability underlying the relative survival method may not be reasonable, it does not have a concerning impact on the estimates of relative survival, as most lung cancer patients die within the first 2 years following diagnosis. This should serve to reassure critics of the use of relative survival when applied to lung cancer data.

British Journal of Cancer advance online publication, 3 May 2012; doi:10.1038/bjc.2012.182 www.bjcancer.com

Summary: the choice between relative and cause-specific survival settings for estimating net survival

- Both aim to estimate the same underlying quantity (net survival).
- Both involve assumptions specific to the approach:

  Cause-specific: Accurate classification of cause-of-death

  Relative: Appropriate estimation of expected survival
- We choose the approach for which we have the strongest belief in the underlying assumptions.
- For population-based studies this is typically relative survival but every study must be evaluated on its specific merits.

Relative survival (estimator) is just one estimator of net survival in a relative survival framework [6]

Table 4. Overview of the two frameworks and measures of cancer patient survival.

	Measure	
	Net survival: competing risks eliminated	Crude survival: in the presence of competing risks
Framework		
Cause-specific: use cause of death information to identify cancer deaths	Cause-specific survival: Censor survival times of noncancer deaths and apply standard estimators such as Kaplan–Meier	Crude probability of death using cause of death: Standard estimators of the cumulative incidence function in the presence of competing risks
Relative survival: contrast all-cause survival of cancer patients to survival of the general population	Net survival: Can be estimated using age-standardized relative survival (Ederer II) or the Pohar Perme estimator of net survival	Crude probability of death in a relative survival framework: Life table approach (Cronin & Feuer) Model-based approach

Many papers compare and discuss the two settings [13]

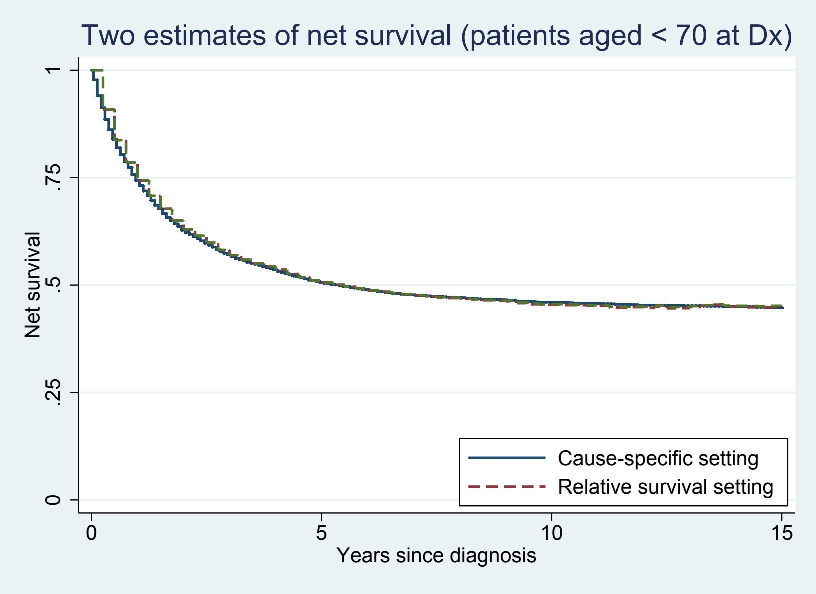
Errors in determination of net survival: cause-specific and relative survival settings

Chloe J. Bright[1], Adam R. Brentnall[2], Kate Wooldrage[3], Jonathon Myles[2], Peter Sasieni[4] and Stephen W. Duffy[2]

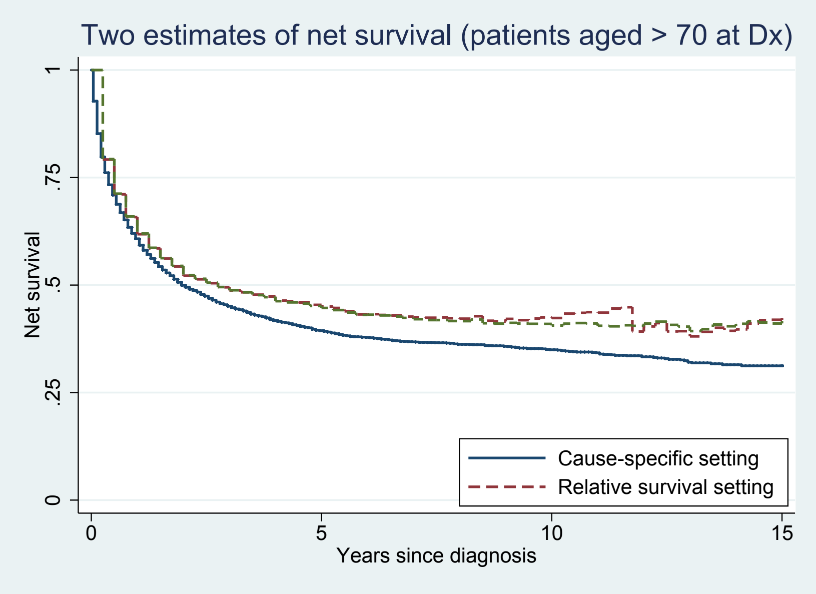
BACKGROUND: Cause-specific and relative survival estimates differ. We aimed to examine these differences in common cancers where by possible identifying the most plausible sources of error in each estimate.
METHODS: Ten-year cause-specific and relative survival were estimated for lung, breast, prostate, ovary, oesophagus and colorectal cancers. The cause-specific survival was corrected for misclassification of cause of death. The Pohar-Perme relative survival estimator was modified by (1) correcting for differences in deaths from ischaemic heart disease (IHD) between cancers and general population; or (2) correcting the population hazard for smoking (lung cancer only).
RESULTS: For all cancers except breast and prostate, relative survival was lower than cause-specific. Correction for published error rates in cause of death gave implausible results. Correction for rates of IHD death gave slightly different relative survival estimates for lung, oesophagus and colorectal cancers. For lung cancer, when the population hazard was inflated for smoking, survival estimates were increased.
CONCLUSION: Results agreed with the consensus that relative survival is usually preferable. However, for some cancers, relative survival might be inaccurate (e.g. lung and prostate). Likely solutions include enhancing life tables to include other demographic variables than age and sex, and to stratify relative survival calculation by cause of death.

British Journal of Cancer (2020) 122:1094–1101; https://doi.org/10.1038/s41416-020-0739-4

Net survival: colon cancer in Finland



Why the difference for older patients?



Cause-specific survival: colon cancer

- Coding of vital status

```
Freq.    Numeric  Label
 4642          0  Alive
 8369          1  Dead: colon cancer
 2549          2  Dead: other
```

- The event of interest is death due to colon cancer.
- Other events are known as 'competing events' or 'competing risks'.
- Based on the research question, we choose between one of two quantities to estimate:
  1. *Eliminate* the competing events (estimate net survival)
  2. *Accommodate* the competing events (estimate crude survival)

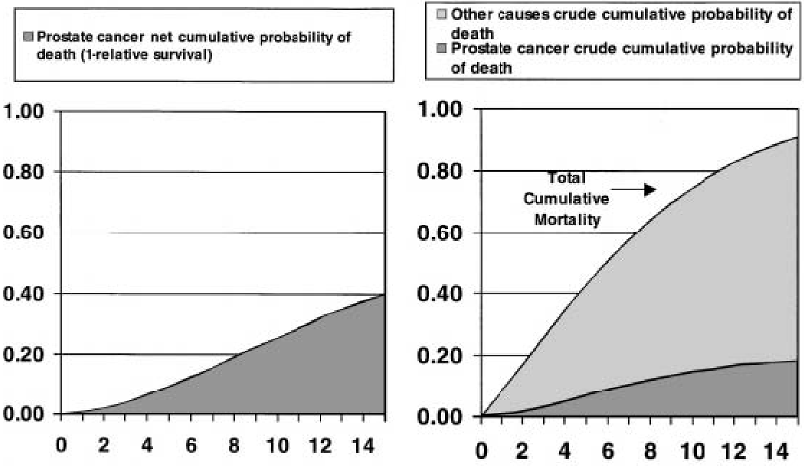
We have a choice of two measures

Net probability of death due to cancer	=	Probability of death in a hypothetical world where the cancer under study is the only possible cause of death

Crude probability of death due to cancer	=	Probability of death in the real world where you may die of other causes before the cancer kills you

- Net probability also known as the marginal probability.
- Crude probability also known as cumulative incidence function.

Net (left) and crude (right) probabilities of death in men with localized prostate cancer aged 70+ at diagnosis (Cronin and Feuer [14])



Explaining net/relative survival to non-scientists

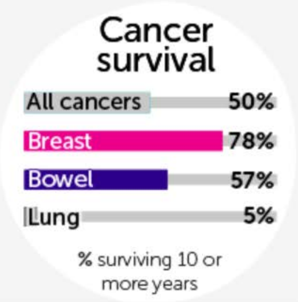
- Organisations that report survival statistics to the general public are often reluctant to describe relative/net survival in a technically correct manner.
- ‘Patients will not understand hypothetical world explanations’ they argue.
- I argue that, if that’s the case, one should report crude (real world) survival rather than estimate net survival and then describe it as something else.

www.cancerresearchuk.org [accessed June 2014] Page has since been updated

10-year net survival was estimated to be 50%.

Cancer survival statistics

- 50% of adult cancer patients diagnosed in 2010-2011 in England and Wales are predicted to survive 10 or more years.
- 46% of men and 54% of women cancer patients diagnosed in 2010-2011 in England and Wales are predicted to survive 10 or more years.
- Cancer survival rates in the UK have doubled in the last 40 years.



www.cancerresearchuk.org/cancer-info/cancerstats/survival/

What does a relative survival of 50% mean? 10-year probabilities of death [15]

Measure	Age 40	Age 60	Age 80
Net prob. of death (1-rel surv)	0.50	0.50	0.50
Crude (actual): cancer death	0.49	0.48	0.42
Crude (actual): non-cancer death	0.02	0.08	0.42
Crude (actual): any cause death	0.51	0.57	0.84

Page has been updated [June 2021]

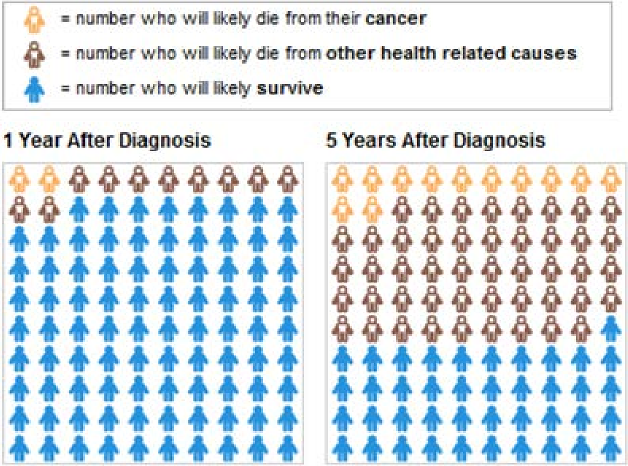
- Same data, new interpretation.
- An improvement, but vague.
- How will readers interpret 'survive cancer'?
- I recognise the need to reduce technical jargon for a general audience.
- Not so for scientific journals.

There is a tradeoff between comparability and interpretability

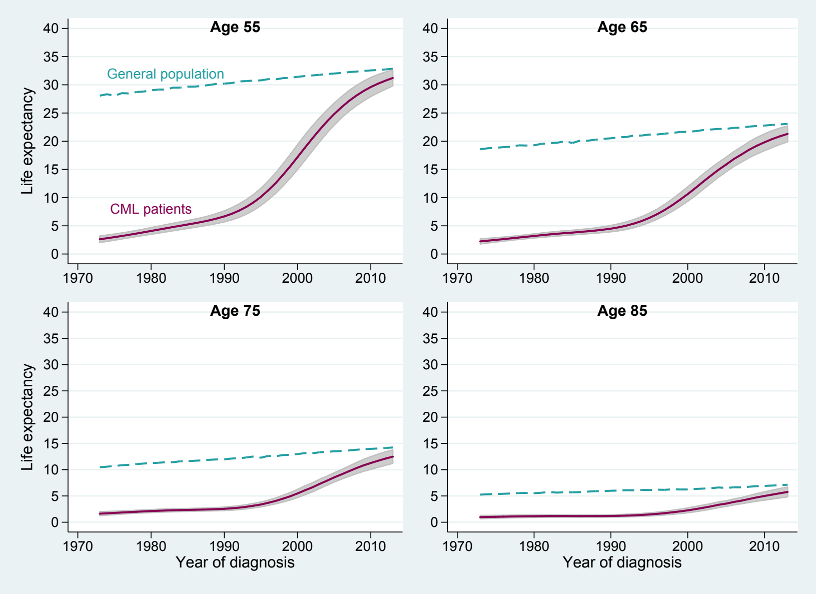
- For international comparisons, we prioritise comparability when choosing a measure.
- Age-standardised net survival is 'the survival one would observe in the hypothetical scenario where cancer were the only possible cause of death and the age distribution of the population were different to what it actually is'.
- This is not directly interpretable in terms of real patients, and we shouldn't try and force an interpretation.

Natural frequencies presented using infographics

'Of 100 patients similar to you, with a cancer similar to that with which you have been diagnosed, we expect ...'



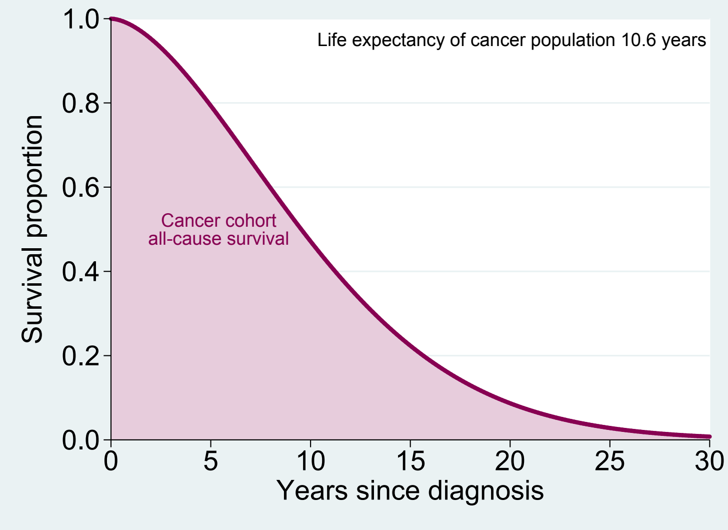
Loss in expectation of life, CML, Sweden



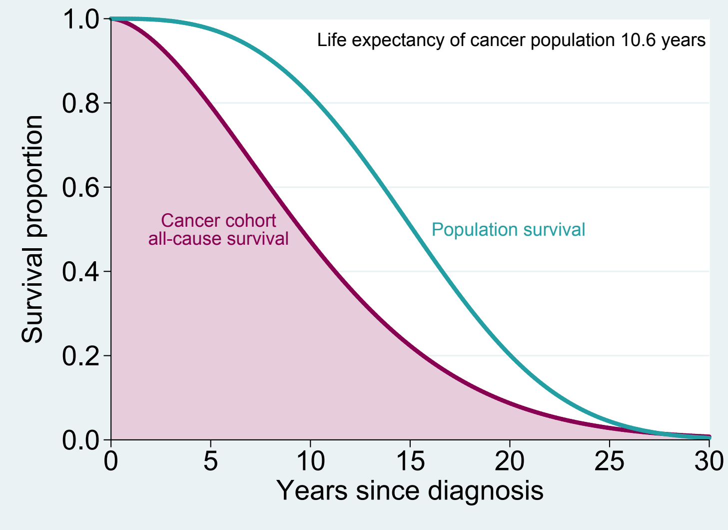
Loss in expectation of life

- A useful summary measure of survival is the mean survival, life expectancy
- The loss in expectation of life is the difference between the mean expected survival (if not diagnosed with cancer) and the mean observed survival (for cancer patients)
- Quantify disease burden in the society "how many life-years are lost due to the disease?"
- Quantify differences between socio-economic groups or countries, "how many life-years are lost in the population due to differences in cancer patient survival between groups?" "how many life-years would be gained if England had the same cancer patient survival as Sweden?"
- Quantify the impact a cancer diagnosis has on a patient's life expectancy

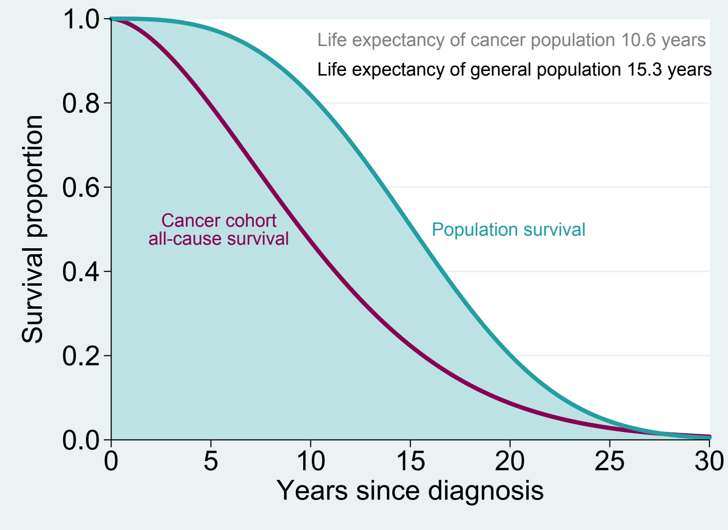
Expectation of life



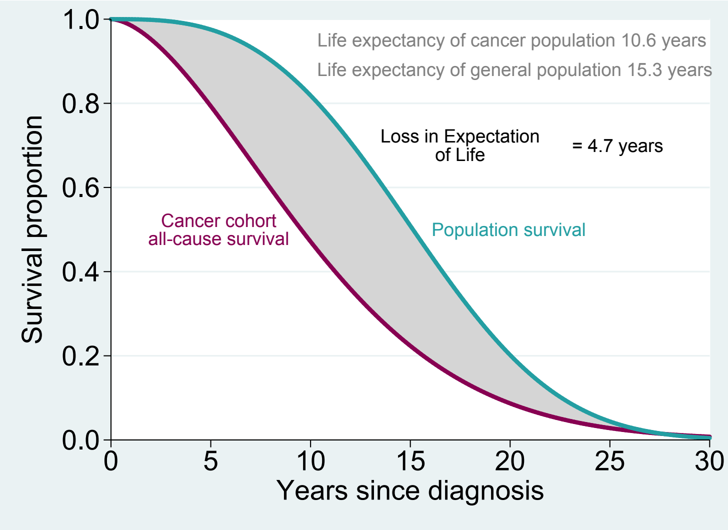
Loss in expectation of life



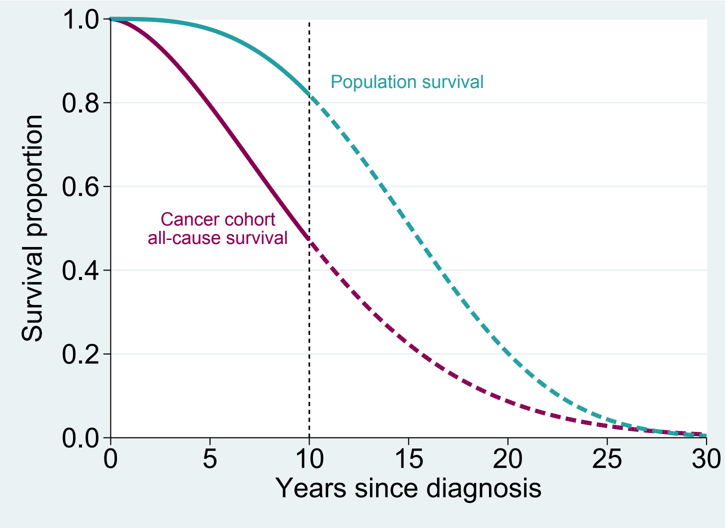
Loss in expectation of life



Loss in expectation of life



Limited follow-up



How do we extrapolate observed survival?

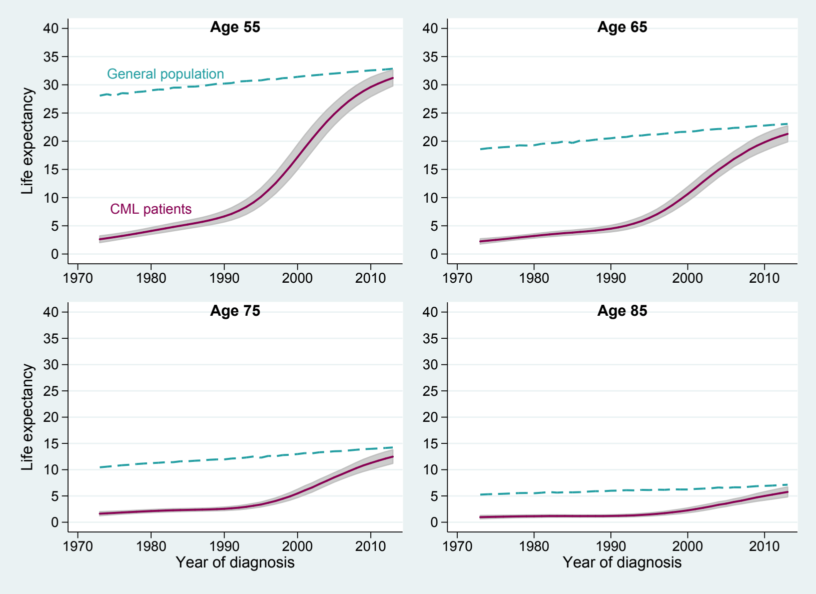
Technical details: recent/current research

- Even though we are now interested in the all-cause survival we will use a relative survival approach

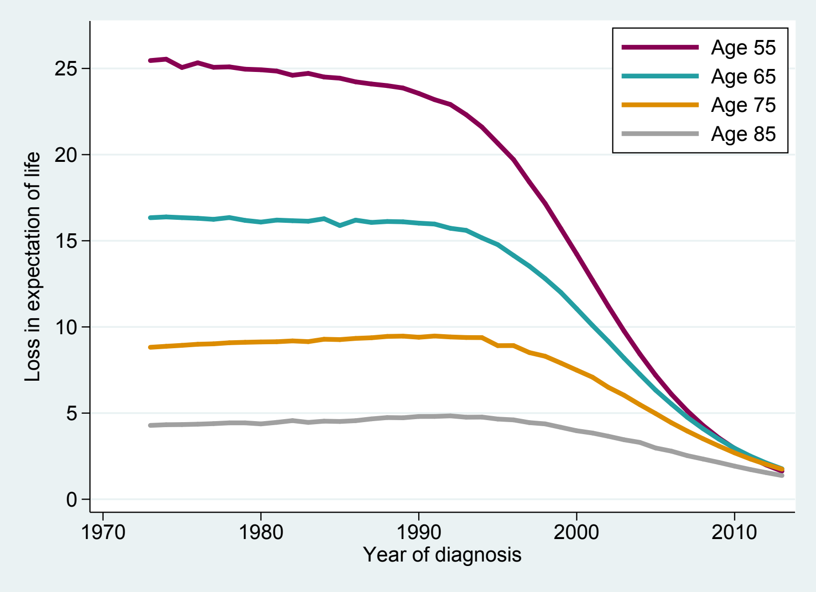
$$S(t) = S^*(t) \times R(t)$$
$$h(t) = h^*(t) + \lambda(t)$$

- Easier to extrapolate $R(t)$ than $S(t)$
- Has been done for grouped data (life tables) [16], by assuming $\lambda(t) = 0$ or $\lambda(t) = c$ after some point in time.
- We estimate in the framework of flexible parametric models [17, 18].

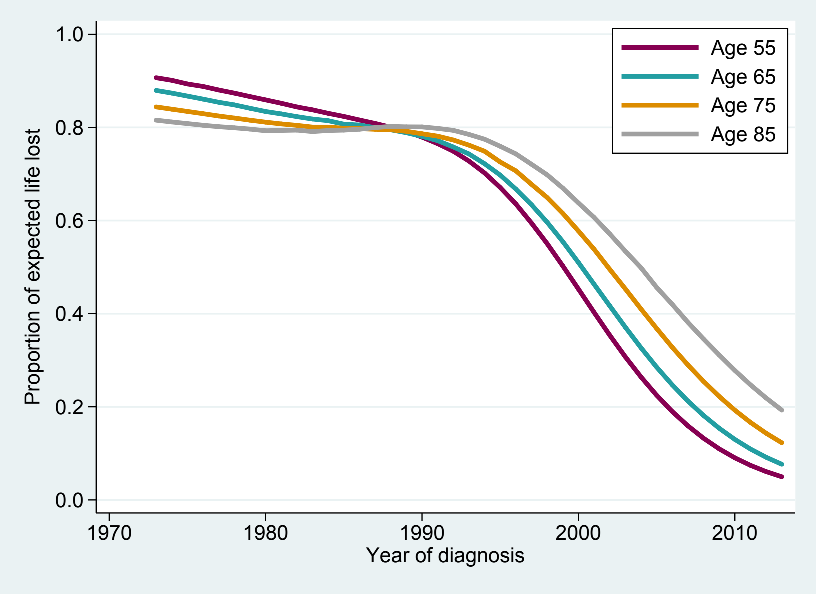
Chronic myeloid leukaemia; Sweden. LE



Chronic myeloid leukaemia; Sweden. LEL



Chronic myeloid leukaemia; Sweden. PELL



References

[1] Adami HO. Sweden: A paradise for epidemiologists? *Lancet* 1996;**347**:588–589.

[2] Dickman PW, Adami HO. Interpreting trends in cancer patient survival. *Journal of Internal Medicine* 2006;**260**:103–117.

[3] Ellis L, Woods LM, Estève J, Eloranta S, Coleman MP, Rachet B. Cancer incidence, survival and mortality: explaining the concepts. *Int J Cancer* 2014;**135**:1774–1782.

[4] Cho H, Mariotto AB, Schwartz LM, Luo J, Woloshin S. When do changes in cancer survival mean progress? The insight from population incidence and mortality. *J Natl Cancer Inst Monogr* 2014;**2014**:187–197.

[5] Eloranta S, Smedby KE, Dickman PW, Andersson TM. Cancer survival statistics for patients and healthcare professionals - a tutorial of real-world data analysis. *Journal of internal medicine* 2021;**289**:12–28.

[6] Kalager M, Adami HO, Lagergren P, Steindorf K, Dickman PW. Cancer outcomes research-a european challenge: measures of the cancer burden. *Molecular oncology* 2021;.

[7] Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, *et al.*. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 2018;.

References 2

[8] Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, *et al.*. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the international cancer benchmarking partnership): an analysis of population-based cancer registry data. *Lancet* 2011;**377**:127–138.

[9] Andersson TL, Rutherford M, Myklebust T, Møller B, Soerjomataram I, Arnold M, *et al.*. Exploring the impact of cancer registry completeness on international cancer survival differences: a simulation study. *British journal of cancer* 2021;**124**:1026–1032.

[10] Cancer Registry of Norway. *Cancer in Norway 2019 - Cancer incidence, mortality, survival and prevalence in Norway*. Oslo: Cancer Registry of Norway, 2020.

[11] Howlader N, Ries LAG, Mariotto AB, Reichman ME, Ruhl J, Cronin KA. Improved estimates of cancer-specific survival rates from population-based data. *J Natl Cancer Inst* 2010;**102**:1584–1598.

[12] Hinchliffe SR, Rutherford MJ, Crowther MJ, Nelson CP, Lambert PC. Should relative survival be used with lung cancer data? *British Journal of Cancer* 2012;**106**:1854–1859.

[13] Bright CJ, Brentnall AR, Wooldrage K, Myles J, Sasieni P, Duffy SW. Errors in determination of net survival: cause-specific and relative survival settings. *British journal of cancer* 2020;**122**:1094–1101.

References 3

[14] Cronin KA, Feuer EJ. Cumulative cause-specific mortality for cancer patients in the presence of other causes: a crude analogue of relative survival. *Statistics in Medicine* 2000;**19**:1729–1740.

[15] Rutherford MJ. Care needed in interpretation of cancer survival measures. *The Lancet* 2014;.

[16] Hakama M, Hakulinen T. Estimating the expectation of life in cancer survival studies with incomplete follow-up information. *Journal of Chronic Diseases* 1977;**30**:585–597.

[17] Andersson TML, Dickman PW, Eloranta S, Lambe M, Lambert PC. Estimating the loss in expectation of life due to cancer using flexible parametric survival models. *Statistics in Medicine* 2013;**32**:5286–5300.

[18] Andersson TML, Dickman PW, Eloranta S, Sjövall A, Lambe M, Lambert PC. The loss in expectation of life after colon cancer: a population-based study. *BMC Cancer* 2015;**15**:412.