European Network of Cancer Registries

Population-based cancer survival metrics workshop

Mark Rutherford

The Joint ENCR-IACR Scientific Conference 2023: Granada, Spain

European Commission

Population-based cancer survival metrics.

Mark J. Rutherford[1,2]

[1]Department of Population Health Sciences,
University of Leicester, UK
[2]Section of Cancer Surveillance, International Agency for Research on Cancer

The Joint ENCR-IACR Scientific Conference 2023: Granada, Spain.
13th November, 2023

Rough Timetable (hopefully time for plenty of discussion, but lots to cover)

09:00–09:05: Welcome

09:05–09:45: Crude and net survival metrics

09:45–10:00: Choice of framework: Relative survival or cause-specific

10:00–10:30: **BREAK**

10:30–10:45: *(continued)* Choice of framework: Relative survival or cause-specific

10:45–11:00: Age-standardisation

11:00–11:20: Survival metrics for different audiences/purposes.

11:20–11:40: Gains in life expectancy and avoidable deaths

11:40–11:55: Discuss estimation approaches available in both a model-based and non-parametric setting

11.55–12.00: Final questions & Closing remarks

Brief background

Population-based cancer survival measures

- Many registries collect information on all cancer patients in a geographically defined area.
- Often sufficient information is collected to calculate survival measures, as well as reporting/monitoring cancer incidence and mortality trends.
- We need:
  - A date of diagnosis (international rules to define this...)
  - A date of death or a date of last follow-up (assume that this is due to a death occurring later than the analysis cut-off date, or a loss to follow-up).
- Many registries assume that they can capture events (deaths) through linkage to a national death register and make the assumption that those without a linked death notification are still alive.
- Alternatively survival-based measures can be ascertained by using active follow-up.

Why survival measures?

- Firstly we want to measure improvements or make comparisons across groups in terms of overall cancer control measures at the population-level.
- This includes early diagnosis, treatment, the functioning of the health system etc. etc.
- We could compare mortality rates... (i.e. the rate of death due to a specific cancer in the population as a whole).
- ... But this will also be influenced by changes in incidence.
- Therefore, it's also interesting to know: are we doing "better" for those that have a diagnosis of cancer. This is where survival-based measures come in.

Interpreting survival, incidence, mortality together...

- We don't have time for masses of detail today, but we always need to consider trends in survival in the context of changes in incidence and any changes (or not) in mortality.
- We need a complete picture to appropriately interpret the overall impact on cancer control.
- Lots of good reference texts on this topic[5, 6], but today we will focus on the survival metrics themselves.

Thoughts for discussion?

1. *What* can we use cancer survival statistics for?
2. *Who* are cancer survival statistics for?

Having thought a little about 1 & 2, we will also come back later to:

3. *Which* measure should we present when?
4. *How* do we direct people to the right statistics for them?

Padlet: Exercise 1

(https://padlet.com/mjr40/cancer_survival_workshop)

1. *What* can we use cancer survival statistics for?
2. *Who* are cancer survival statistics for?

Aims of the analysis?

1. General reporting of cancer survival metrics as descriptive information - cancer registry reports, patient/carer information websites/charity pages?
2. Comparative studies of cancer survival across population groups - e.g. international comparisons[4, 33, 39], comparison across population subgroups in national analyses (e.g. socioeconomic[34], sex[37], race[38], calendar time, etc.).
3. Causal comparison of intervening on cancer survival differences - what if...?

Available metrics - we'll come back to these.

1. Net survival
2. All-cause survival
3. Crude probabilities of death due to cancer/other causes.
4. Life expectancy
5. Conditional version of above metrics (i.e. conditioning on surviving X years after diagnosis)
6. ...

A couple of papers with discussion of these metrics and their potential differential focus for the audience: [1, 2]

Further background - survival measures

- We often want to make fair comparisons of prognosis across population groups.
- Ideally, we don't want differences in risks of other outcomes impacting on our metrics for our disease of interest (cancer).
- Often population groups who are unequal in terms of cancer survival also have different competing risks due to other causes (e.g. socioeconomic groups).
- Groups could be age, socioeconomic class, countries, calendar periods...

Prognosis across groups

Why might there be differences in all-cause survival between two groups?

1. Differences in disease-specific mortality rates.
2. Differences in other-cause mortality rates.
3. Differences in age (or other covariate) distribution.

Prognosis across groups - How do we typically solve them now?

Why might there be differences in all-cause survival between two groups?

1. Differences in disease-specific mortality rates.
   - We tend to focus on this - calculate metrics depending on cancer-specific mortality differences.
2. Differences in other-cause mortality rates.
3. Differences in age (or other covariate) distribution.

Prognosis across groups - How do we typically solve them now?

Why might there be differences in all-cause survival between two groups?

1. Differences in disease-specific mortality rates.
   - We tend to focus on this - calculate metrics depending on cancer-specific mortality differences.
2. Differences in other-cause mortality rates.
   - We often try to eradicate other-cause mortality differences (net measures).
3. Differences in age (or other covariate) distribution.

Prognosis across groups - How do we typically solve them now?

Why might there be differences in all-cause survival between two groups?

1. Differences in disease-specific mortality rates.
   - We tend to focus on this - calculate metrics depending on cancer-specific mortality differences.
2. Differences in other-cause mortality rates.
   - We often try to eradicate other-cause mortality differences (net measures).
3. Differences in age (or other covariate) distribution.
   - We standardise to some external age-standard, which may be far from our population age distribution.

Prognosis across groups - How do we typically solve them now?

Why might there be differences in all-cause survival between two groups?

1. Differences in disease-specific mortality rates.
   - We tend to focus on this - calculate metrics depending on cancer-specific mortality differences.
2. Differences in other-cause mortality rates.
   - We often try to eradicate other-cause mortality differences (net measures).
3. Differences in age (or other covariate) distribution.
   - We standardise to some external age-standard, which may be far from our population.

This makes pretty hypothetical metrics. Should we do better for some/all purposes? These net, standardised metrics are also often accessible on more patient-focussed material...

Motivation/Background

- So, we *often* want group comparisons and general reporting of survival metrics to be independent of different competing causes of death across groups.
- Therefore, we try to isolate the impact of cancer alone on mortality (and base our metrics around that).
- We essentially isolate cancer-specific mortality - but people prefer survival-based metrics...
- Hence, we often end up with relative/cause-specific/net survival measures.

Crude and net survival metrics

Introduction

- Cancer survival measures generally are reported as net measures, where deaths from other causes are eliminated (accounted for) in the estimation in either a cause-specific or relative survival framework.
- These estimates are typically presented as population summary (average) and also age-standardised.
- These estimates are undoubtedly useful for comparability, but are perhaps overused.
- Some thought should be given to the purpose and audience of cancer survival measures; this may alter what we would choose to present.
- The standard approach of estimating net survival is useful for comparing populations, but not necessarily relevant to individual patients...

What is net survival?

- Net survival measures the survival experience of patients in the world where only mortality due to the cancer of interest can act upon patients. These metrics are under the hypothetical scenario where cancer is the only possible cause of death.
- It is not possible to observe this net measure in the real-world where patients can die of any number of causes.
- We therefore use approaches in a relative survival or cause-specific survival framework to attempt to estimate net survival under certain unverifiable assumptions (synonymous with arguments and estimation in competing risks).

What are competing events?

Events that prevent the occurrence of the event of interest i.e. dying from one cause means that the time-to-death for the competing cause is never observed.

- How to deal with competing events, depends on what the research questions is.
- Direct effect on the event of interest eliminating the competing events (net setting)?
- Effect on the event of interest accommodating the competing events (real-world setting)?

Measures that eliminate competing events

- Net survival: when we are interested in cancer-specific mortality.
- We ignore deaths from other causes to estimate net survival - hypothetical world where cancer is only cause of death.
- This helps comparability across population subgroups with differential background mortality - we remove distortions from other-cause mortality differences.
- Net survival can be estimated using:
  - Cause-specific survival using cause of death information
  - Relative survival using available population lifetables

UNDER ASSUMPTIONS!

Assumptions for both frameworks

Relative Survival :

1. Other than the set of measured covariates in our lifetable and analysis, no other factor should affect both the time-to-death due to the disease of interest and time-to-death due to other causes.
2. Appropriate expected mortality information. This means that the mortality rate due to other causes for the cancer patients is the same as that in the population lifetable.

Cause-specific Survival :

1. Other than the set of measured covariates in our analysis, no other factor should affect both the time-to-death due to the disease of interest and time-to-death due to other causes.
2. Appropriate classification of cause of death information. The ability to correctly ascertain if a specific individual has died due to cancer or due to another cause.

What are crude measures? - Why do we swap to crude probabilities of death?

- Crude measures entail no elimination of competing events. They refer to a setting where competing events are present.
- Useful for understanding the anticipated prognosis of patients, for risk communication and healthcare planning.
- Because now at least two causes of death exists - we tend to swap to the mortality scale...
- For instance, now to "survive from cancer" - you could either still be alive, or have died due to something else - awkward interpretation.
- We therefore break-down the all-cause probability of death into component parts - that due to cancer and that due to other causes.
- However, the probability of dying due to cancer will be influenced by the "weight" of mortality due to other causes - need to consider this when comparing across groups.

Crude probabilities of death

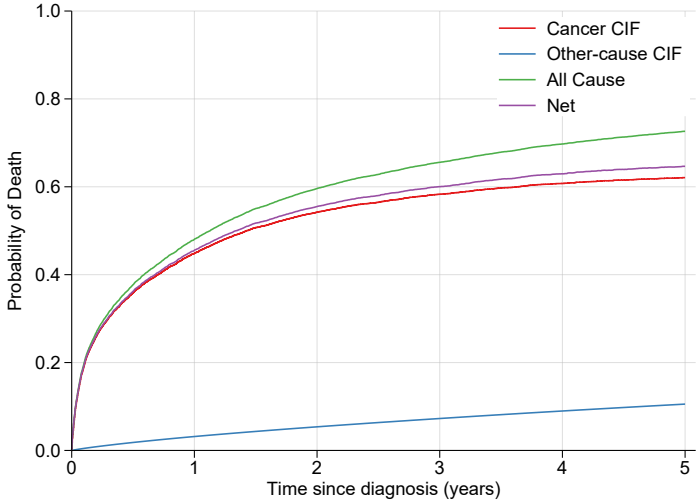
The total marginal all-cause hazard rate can be partitioned (in age-groups or in the population overall) into component parts, depending on the competing causes.
We then construct the cumulative incidence functions for deaths due to cancer, $F_c(t)$ and deaths due to other causes $F_o(t)$ as follows:

$$F_c(t) = \int_0^t S(u)h_c(u)du$$

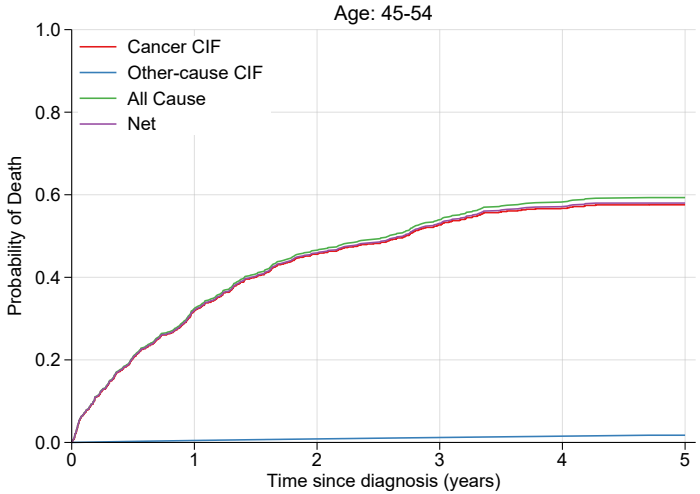
$$F_o(t) = \int_0^t S(u)h_o(u)du$$

where $S(t)$ is the all-cause survival function. We can also estimates these quantities from a relative survival framework[3] - rather than a cause-specific framework. We will come to this.

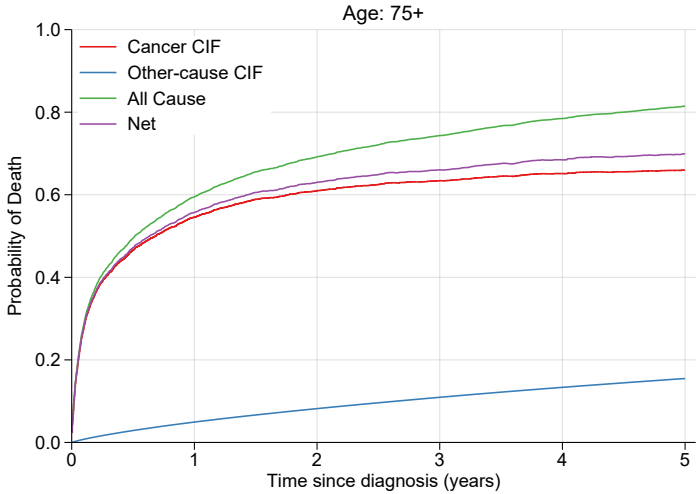
Net vs Crude - Colon, England (1985-1990) : All-ages



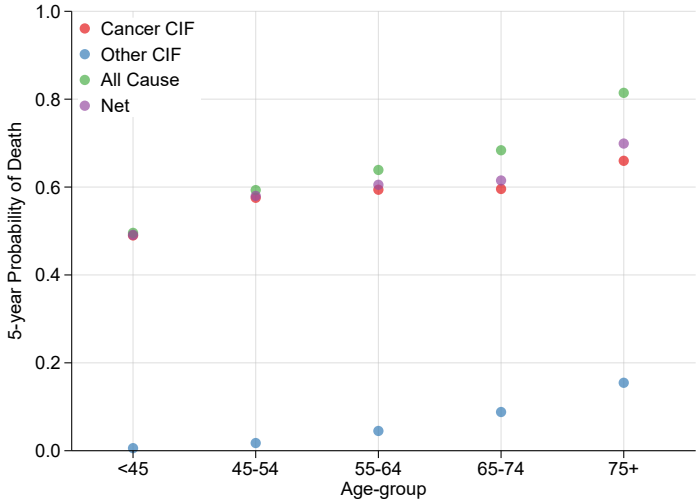
Net vs Crude - Colon, England (1985-1990) : Age 45-54



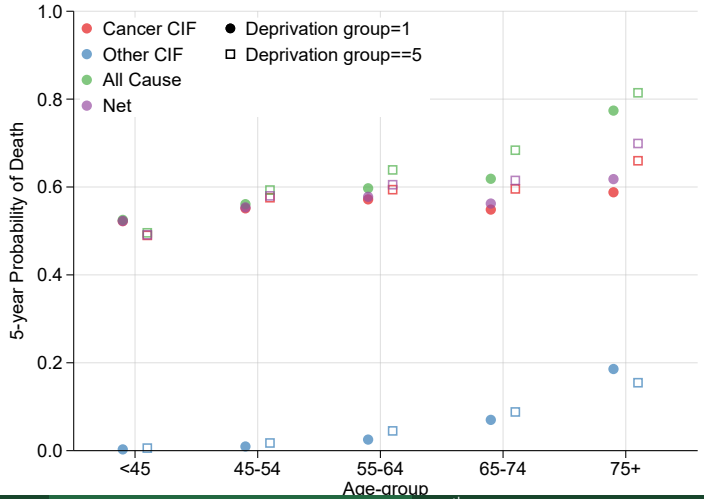
Net vs Crude - Colon, England (1985-1990) : Age 75+



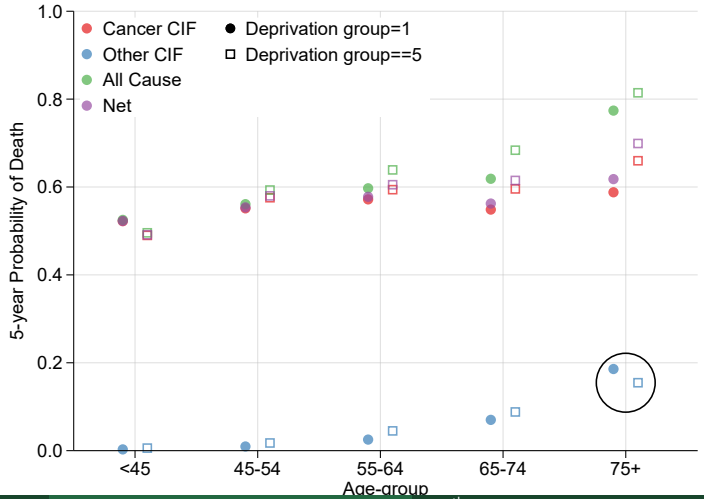
Net vs Crude - Colon, England (1985-1990) : Age-group estimates



Net vs Crude - Colon, England (1985-1990) : Age-group estimates by deprivation



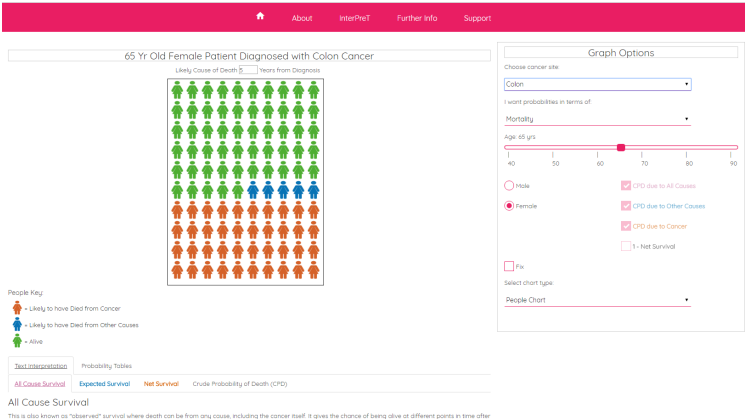
Net vs Crude - Colon, England (1985-1990) : Age-group estimates by deprivation



InterPreT: We will look at a demo next

An Interactive cancer survival Prediction Tool

`https://interpret.le.ac.uk/` [Accessed 1st November, 2023].



Positives and negatives...crude vs net metrics

Positives	**Crude**	**Net**
	Ease of interpretation	Comparability across groups
	Real world measures	
	...	...
Negatives	**Crude**	**Net**
	Lack comparability?	Hypothetical world
	Depend on competing risk	Hard to interpret?
	...	...

Suggestions for more elements for this table?

Examples of crude probability reporting and metrics

- There have been many examples of reporting and encouraging the reporting of crude probability measures alongside or to supplement net measures [35, 9, 12, 13].
- The methods papers for undertaking these analysis in a relative survival context have been available for many years[3].
- The uptake of these metrics is somewhat hindered perhaps by a favouring for comparability over interpretability?
- Another feature that we haven't discussed much is the consideration for both net[40] and crude measures[9] to use conditional metrics.
- Conditional metrics can be a powerful way to give updated information for individuals that have already survived for a fixed period beyond diagnosis.

Reference adjusted metrics - the best of both worlds? See Paul Lambert's talk tomorrow...

- The key to reference adjustment approaches is to keep the "fairness" of net metrics when comparing across population group, but to convert back to a real-world "crude" format.
- To do so requires a choice of a reference standard for the expected mortality rates (in the relative survival framework).
- Paul Lambert will discuss this for international comparison of survival metrics tomorrow.
- Further reading... [10, 29, 31, 32]

Choice of framework: Relative survival or cause-specific

Net measures or crude measures, we still need to pick an estimation framework...

- Whichever choice of metric we go for, there's a choice to be made in estimation framework.
- We can choose a relative survival framework - where the key input will be the population lifetable we contrast to.
- Alternatively, we can use cause of death information to ascertain which deaths are directly due to cancer and which are not. This will then be the cause-specific framework.

Choice between relative and cause-specific frameworks for estimating net survival

- Both frameworks require the independence assumption.
- Each framework requires a specific assumption:

  Cause-specific Accurate classification of cause-of-death

  Relative Appropriate estimation of expected survival
- We choose the framework for which we have the strongest belief in the underlying assumptions.
- For population-based studies it has become quite common to use the relative survival framework but every study must be evaluated on its specific merits.

Why excess mortality/relative survival?
(I) We are interested in cancer-specific mortality.

- We end up using (more often than not) using excess mortality rather than cause-specific mortality.
- We split the total mortality (hazard), $h_i(t)$, into component parts; that due to background mortality, $h_i^*(t)$, and the excess due to the disease, $\lambda_i(t)$.

$$h_i(t) = h_i^*(t) + \lambda_i(t)$$

- **REASON 1:** We wish to focus on the mortality due to cancer alone.

Why excess mortality?
(II) We don't trust/have cause of death information

- Cause of death information has been shown to be unreliable, particularly for certain population groups (the elderly, for instance).
- This information may also be completely unavailable in certain settings.
- Complications of cause of death classification can also arise - does a death that is directly related to surgery get appropriately coded as caused by cancer etc.?

Excess mortality/relative survival

We split the total hazard, $h_i(t)$, into component parts; that due to background mortality, $h_i^*(t)$, and the excess due to the disease, $\lambda_i(t)$.

$$h_i(t) = h_i^*(t) + \lambda_i(t) \tag{1}$$

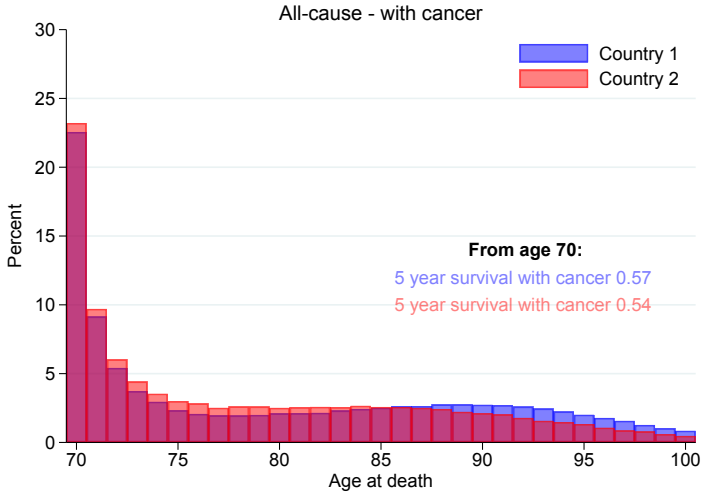
We convert to the survival scale:

$$S_i(t) = S_i^*(t) R_i(t) \tag{2}$$

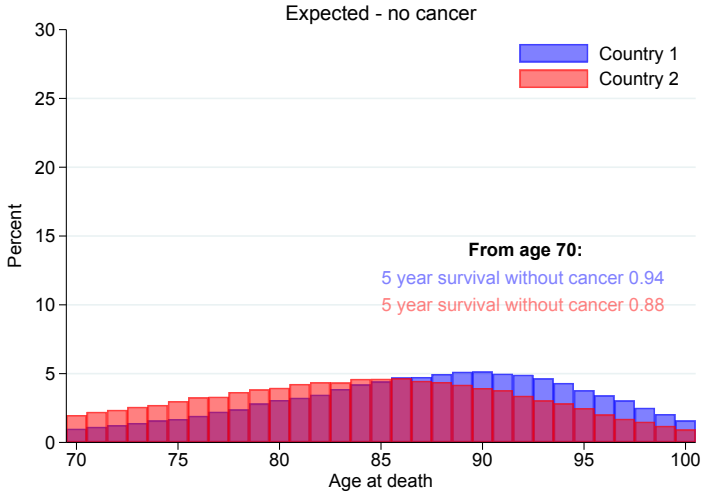
And see why it's called relative survival:

$$R_i(t) = \frac{S_i(t)}{S_i^*(t)} \tag{3}$$

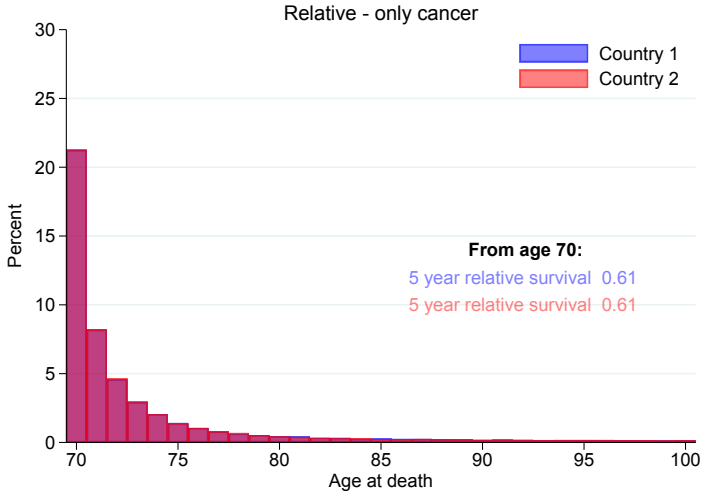
Why excess mortality/relative survival? Why net measures?



Why excess mortality/relative survival? Why net measures?



Why excess mortality/relative survival? Why net measures?



Relative survival vs cause-specific framework

- There are many papers discussing the merits of relative survival vs cause-specific approach, and drawing comparisons [14, 16, 17, 18, 19, 20, 23, 22, 24].
- A key example where the assumptions of relative survival may be unreasonable is lung cancer. In that case, a larger proportion of our cancer cases will be smokers, but we are not reflecting that in our choice of population lifetable.
- There are also a number of approaches that have tried to make sensible adjustments to cause of death classification with the purpose of using that for cause-specific survival - particularly in the US[36]. Avenues for this are also discussed here in a UK paper[21].
- We wrote a recent paper trying to make sure that people are being fair when comparing across the frameworks[15], which is also covered in earlier work[22].

Further thoughts

- With more complex analyses being undertaken, it may be difficult to create the appropriate lifetables with sufficient granularity (e.g. analyses stratifying by comorbidity status).
- From within both frameworks, we still have a range of choices for the metrics to then go on to present.

Age-standardisation

Age-standardisation

- Standardisation is very common in epidemiology to try and ensure comparisons are fair.
- When making comparisons between groups we should compare "like-with-like".
- When we compare incidence and mortality between different populations we always need to think about adjusting for age (and other key covariate) differences.
- Thus we usually use age-standardised estimates when presenting incidence and mortality. The same ideas carry over to (relative) survival.

Traditional Age standardisation

- Relative survival is estimated separately in each of S age groups.
- The age standardised estimate is a weighted average of the relative survival in each age group, $R_j(t)$.

$$R_s(t) = \sum_{j=1}^{S} w_j R_j(t)$$

- The weights, w_j could be based on age distribution observed in the study (internal age standardisation) - or an external standard.
- It is important to realise that there may be huge variation in relative survival between age groups, but this can be 'lost' when only presenting age standardised estimates.

External Age standardisation

- The main reason to externally standardise is that we want to compare relative survival between different groups which may have a different age distribution.
- In doing this we are forcing a different age distribution onto our study population to that they actually have.
- This means that we are estimating survival in a hypothetical world where you can only die of the cancer under study **and** if the population had a different age distribution to what they actually have!
- We should be very cautious about putting a real world interpretation on this and remember that we are standardising in order to make fair comparisons.

External Age standardisation 2

- The analyst has the choice of what age distribution to use. This could be:
  - An agreed standard age distribution (See following slide)
  - The age distribution in a particular calendar period when comparing survival between different calendar periods.
  - The age distribution in a particular subgroup.

International Cancer Survival Standard weights

The three International Cancer Survival Standard (ICSS) weights used for age-standardisation of relative survival[27]

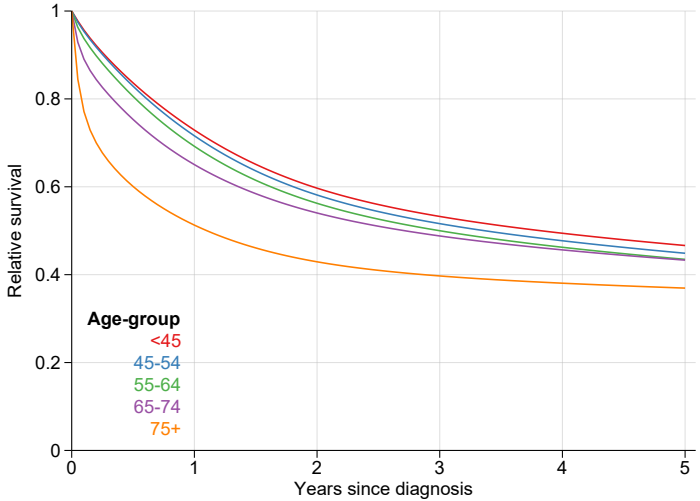
Age	ICSS 1[a]	ICSS 2[b]	ICSS 3[c]
15-44 years	0.07	0.28	0.60
45-54 years	0.12	0.17	0.10
55-64 years	0.23	0.21	0.10
65-74 years	0.29	0.20	0.10
75+ years	0.29	0.14	0.10

[a] Lip, tongue, salivary glands, oral cavity, oropharynx, hypopharynx, head and neck, oesophagus, stomach, small intestine, colon, rectum, liver, biliary tract, pancreas, nasal cavities, larynx, lung, pleura, breast, corpus uteri, ovary, vagina and vulva, penis, bladder, kidney, choroid melanoma, non-Hodgkin lymphoma, multiple myeloma, chronic lymphatic leukaemia, acute myeloid leukaemia, chronic myeloid leukaemia, leukaemia, prostate

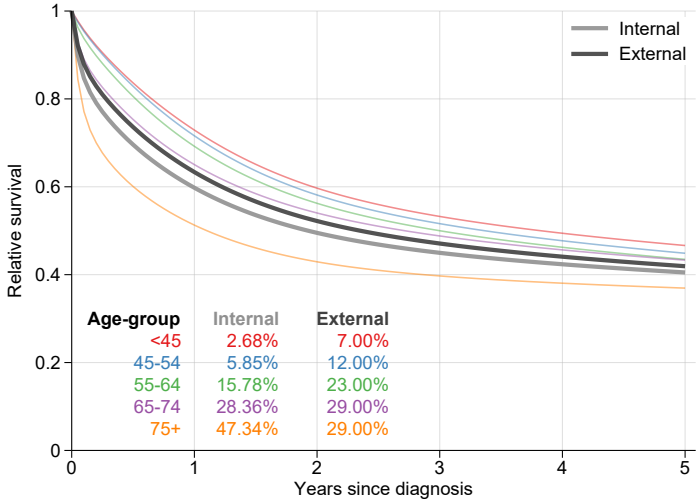
[b] Nasopharynx, soft tissues, melanoma, cervix uteri, brain, thyroid gland, bone

[c] Testis, Hodgkin lymphoma, acute lymphatic leukaemia

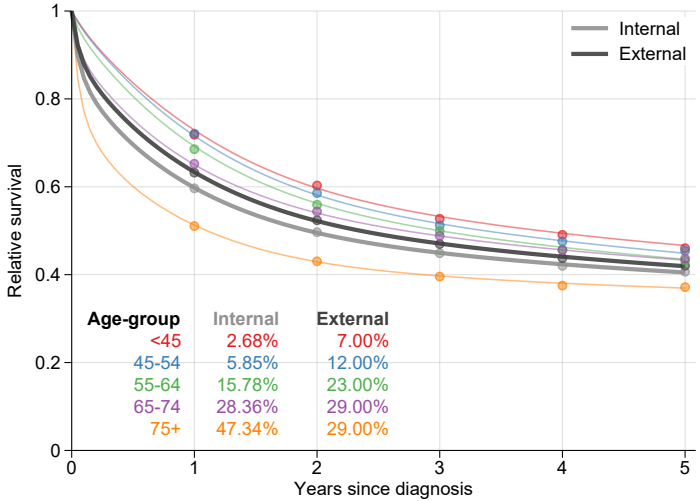
Age-standardisation example - Colon, England (1985-1990)



Age-standardisation example - Colon, England (1985-1990)



Age-standardisation example - Colon, England (1985-1990)



Brenner et al. (2004) 'alternative approach'[28]

- Aim is to obtain age standardised estimate without having to stratify analysis by age groups.
- Weights are individually assigned to all patients depending on their age-group.
- If a patient has weight 1.8 then this patient contributes 1.8 units to the 'at risk' column at entry and 1.8 units to the deaths column at death (or the withdrawal column at censoring).
- Weights are higher than 1 in age groups under-represented in the study population compared with the standard population and vice versa.
- Can be used with Pohar Perme or a model-based approach.
- **Advantage:** we don't require age-specific estimates in each stratum. The method can be useful with sparse data.

Summary & Cautions around age standardisation?

- The concept of age-standardization is fairly simple - it is just a weighted average of different relative survival estimates.
- However, there is often large variation in relative survival by age - collapsing to an average summary measure masks this variation.
- There may also be interesting differences across compared population groups by age too - i.e. more stark inequalities in survival at older ages.
- Using ICSS weights is a good option for international comparison studies - but these can often be quite far from the internal age distribution in the sample (as we saw in the example).
- At least some thought should be given on the purpose of the analysis again when choosing the age standard.
- Similar standardisation could also be done for crude metrics etc. - but again thought is needed on the right choice.

Survival metrics for different audiences/purposes

What more *could* we do?

- Firstly, we *could* create more individualised predictions by using statistical models; even by using only the most basic covariate information such as age, sex, stage etc.
- Secondly, we *could* present real-world estimates rather than net measures, so that people can appreciate their true risk of being alive X years down the line following a diagnosis of cancer.
- Finally, we *could* also consider using different metrics and methods of presentation in order to make the information easier to understand and interpret.

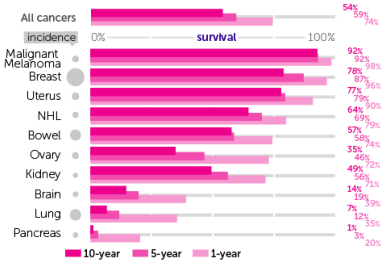
Alternative Metrics

- Metrics to explain impact of cancer:
  - Loss in life expectancy.
  - Centiles of all-cause survival distribution.
  - Conditional survival.
  - Crude probabilities of death.
- Metrics to show inequalities:
  - Difference in total deaths (avoidable/preventable deaths).
  - Gain in life expectancy.

Choose a measure, framework, and estimator

CRUK website - Top 10 Cancer (Female) Link

Age-Standardised One-, Five- and Ten-Year Net Survival, Selected Cancers, Adults (Aged 15-99), England and Wales, 2010-2011



Five- and ten-year survival for 2010-2011 is predicted using an excess hazard statistical model.
Survival for bowel cancer is a weighted average derived from data for colon (C18) and rectum cancer (C19-C20, C21.8)

Source: cruk.org/cancerstats

CANCER RESEARCH UK

CRUK website - Bowel Cancer link

Survival for all stages of bowel cancer

Generally for people with bowel cancer in England:

- almost 80 out of 100 people (almost 80%) survive their cancer for 1 year or more
- almost 60 out of 100 people (almost 60%) survive their cancer for 5 years or more
- almost 55 out of 100 people (almost 55%) survive their cancer for 10 years or more

What does a net survival of 50% mean?
10-year probabilities of death [41]

Measure	Age 40	Age 60	Age 80
Net prob. of death (1-net 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

Cancer in Norway Special Issue 2021 Link to Special Issue

- Weblink above to a Special Issue report for the Cancer in Norway publication.
- Uses a broader range of metrics, with a view to reporting to different audiences.
- Excellent introduction section written by Paul Dickman on many of the issues we have been discussing today.

Padlet: Exercise 2

(https://padlet.com/mjr40/cancer_survival_workshop_2)

3. *Which* measure should we present when?
4. *How* do we direct people to the right statistics for them?

Gains in life expectancy and Avoidable Deaths

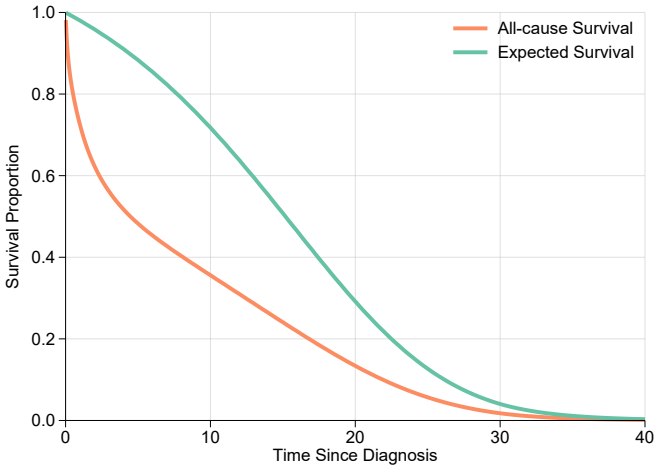
Metrics for direct comparison across groups in the real-world...

- Potential gains in life expectancy and the potential number of deaths that could be avoided have been used to better quantify population inequalities in cancer survival.
- These are real-world metrics.
- But the key is we first isolate the net survival differences across groups - we then convert back to the real-world using the other-cause survival experience of a single group of interest.
- We ask What if? we could remove inequalities in cancer survival but we kept other cause mortality to be the same for the disadvantaged group.

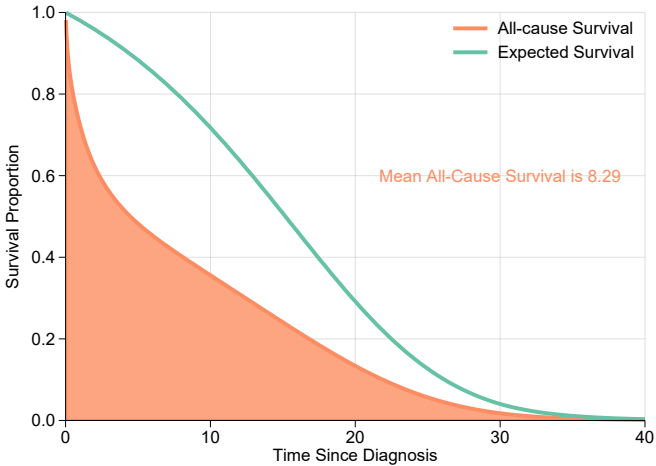
Life Expectancy and Loss of Life

- We can calculate life expectancy based on key characteristics (e.g. age, sex, socioeconomic status etc. etc.)
- Having a diagnosis of cancer is known to reduce life expectancy.
- We can calculate the average loss in life expectancy for patient groups to quantify the impact of cancer.
- This measure is age-dependent; younger patients have more life to lose.
- We can estimate this quantity from *exactly* the same statistical model as used for relative survival analyses.

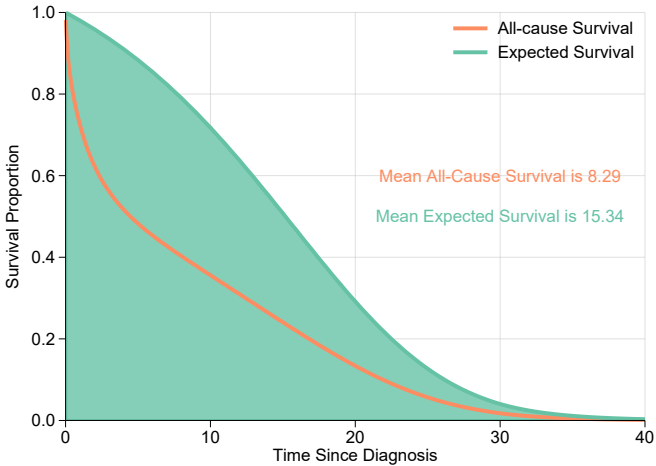
Average Loss in Expectation of Life (Colon, age 70, females)



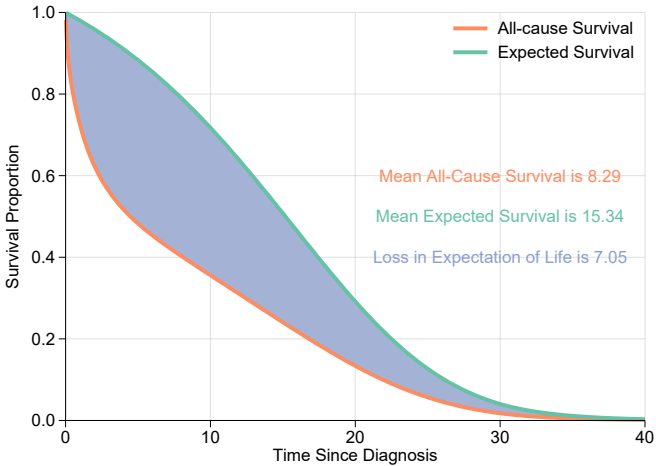
Average Loss in Expectation of Life (Colon, age 70, females)



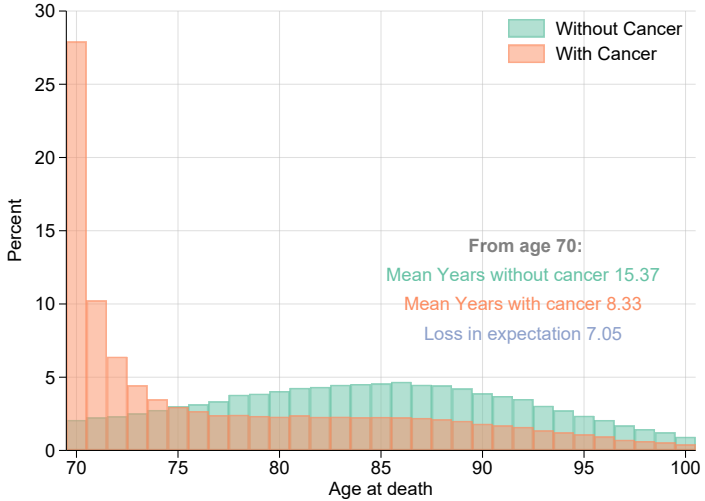
Average Loss in Expectation of Life (Colon, age 70, females)



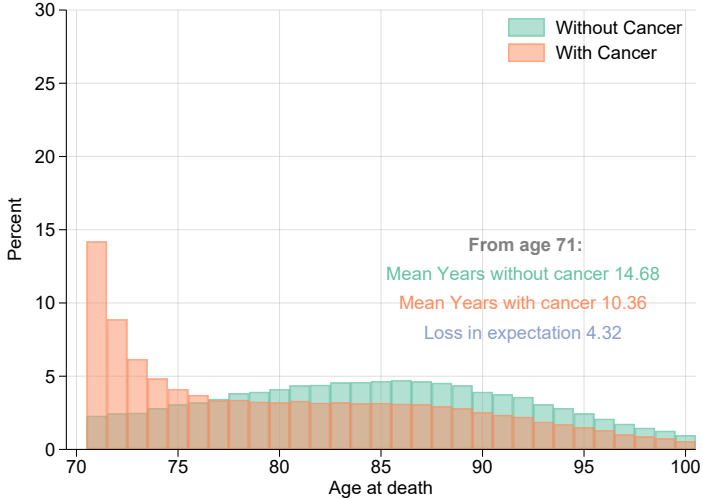
Average Loss in Expectation of Life (Colon, age 70, females)



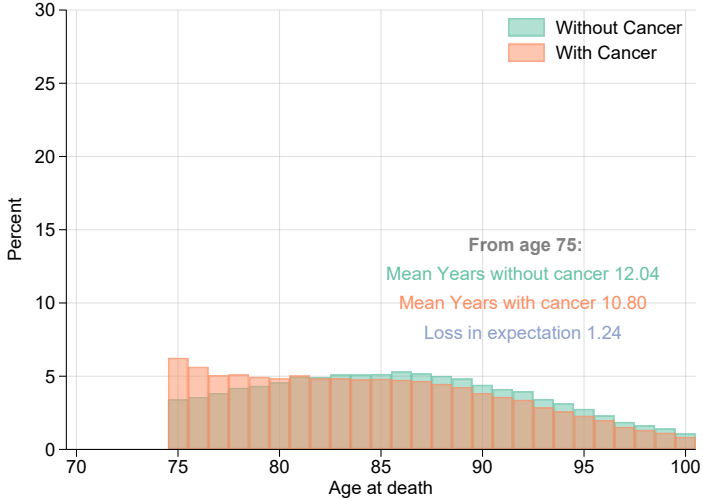
Loss in Expectation of Life



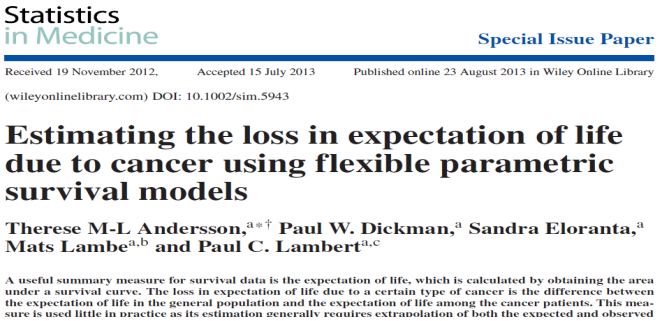
Conditional Loss in Expectation of Life (Conditional on 1 years)



Conditional Loss in Expectation of Life (Conditional on 5 years)



Andersson Paper



Statistics
in Medicine

Special Issue Paper

Received 19 November 2012, Accepted 15 July 2013 Published online 23 August 2013 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5943

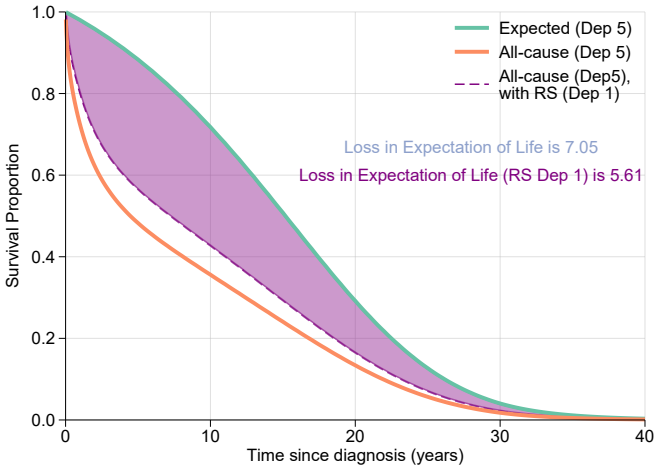
Estimating the loss in expectation of life due to cancer using flexible parametric survival models

Therese M-L Andersson,[a*†] Paul W. Dickman,[a] Sandra Eloranta,[a] Mats Lambe[a,b] and Paul C. Lambert[a,c]

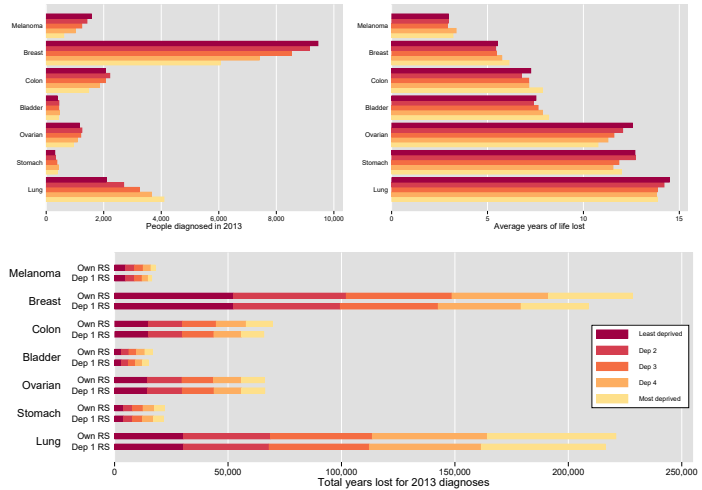
A useful summary measure for survival data is the expectation of life, which is calculated by obtaining the area under a survival curve. The loss in expectation of life due to a certain type of cancer is the difference between the expectation of life in the general population and the expectation of life among the cancer patients. This measure is used little in practice as its estimation generally requires extrapolation of both the expected and observed

- Project forward relative survival using a linear constraint.
- Gives good approximation to all-cause survival observed in practice.

Impact of inequalities:
Colon Females; Deprivation group 5, age 70



Gains in expectation of life? (females only): removing inequities?[34]



Avoidable deaths

- A similar approach to quantifying the value of removing inequalities can be used to report the number of avoidable deaths at a particular point in time.
- This uses almost exactly the same "What if"? approach, but doesn't extrapolate to a lifetime horizon as we are doing in the life expectancy calculation.
- This is becoming an increasingly popular approach for exploring inequalities. Here are some example studies: [11, 42, 43].
- More complex analyses have begun to unpick the reason for the inequalities - such as differences in stage at diagnosis across the compared groups e.g. [44, 45].

Summary

- For these analyses we ask: "What if?" we could remove inequalities in net survival but we kept other cause mortality to be the same for the disadvantaged group.
- These are nice, real-world summaries of the impact of removing inequities in net survival across population groups.
- This is strongly related to reference adjusted survival metrics that I mentioned earlier.

Model-based vs non-parametric, and estimation approaches

Choice of approach

- So far, we've discussed mostly the concepts of net measures, crude probabilities, life expectancy etc.
- Many of the metrics we have discussed can be estimated non-parametrically.
- Some of the more advanced metrics lend themselves to a model-based framework.
- This is also true if interest is in variation across multiple (including continuous) covariates.
- The nice feature of the model-based approach is that we can make predictions both for a specific covariate profile (conditional measures), and we can collapse back to average (marginal) measures as a population summary.

Non-parametric approaches: relative survival framework

- Pohar Perme approach[26] favoured in the relative survival framework for net survival estimation.
- Cronin and Feuer[3] described an approach for estimating crude probabilities non-parametrically in the relative survival framework.

Model-based approaches: relative survival framework

- We often use flexible parametric excess mortality models [46, 47] to estimate the range of metrics we have discussed today.
- Other modelling approaches exist too[7].
- I'm concentrating on the relative survival framework, but more standard regression approaches can be used in the cause-specific setting (e.g. cause-specific Kaplan-Meier (weighted - see [15]), a Cox model, Fine & Gray model etc.)

Checking key modelling assumptions

- If we model, then we have to make modelling assumptions (e.g. proportional excess hazards, functional form of covariate effects, to include interactions or not).
- It's key that we use the non-parametric estimates as an approach to check if our modelling assumptions are reasonable.

Software implementations

Stata :

- `stpp` : Pohar Perme estimates, crude probabilities in continuous time.
- `strs` : Life-table approximation: non-parametric relative survival, crude probabilities.
- `stpm3` : model-based implementation for flexible parametric modelling.

R :

- Lots available in the `relsurv` package for non-parametric estimates[25]
- Model-based options for flexible parametric modelling: `rstpm2` package, `flexsurv` package[50], `mexhaz` package[49] more...

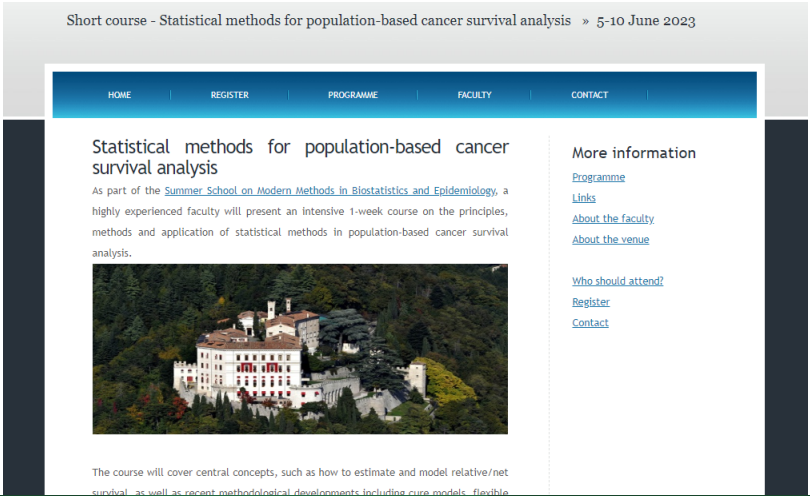
SAS :

- See SAS macros from Ron Dewar `https://github.com/FlexSurv/repo`

SEER*Stat : Many estimation approaches available in SEER∗Stat.

A week-long course on this and more...

http://cansurv.net/index.html



Final thoughts

Final thoughts

- More thoughts should be given to the audience of the survival metrics we produce.
- This may govern the metrics we choose, and how we (age-)standardise our metrics.
- Be careful to consider the estimation framework (cause-specific vs relative).
- Should we stop reporting net metrics all together?

Questions/Thoughts?

`https://github.com/MJRutherford9/cancer_survival_workshop`

E-mail: mark.rutherford@le.ac.uk , **Twitter/X:** MJRutherford9

Selected References/Further Reading

Selected References

[1] 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.

[2] Belot A, Ndiaye A, Luque-Fernandez MA, Kipourou DK, Maringe C, Rubio FJ, Rachet B. Summarizing and communicating on survival data according to the audience: a tutorial on different measures illustrated with population-based cancer registry data. Clinical epidemiology 2019;**11**:53–65.

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

[4] Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić 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 London England 2018;**391**:1023–1075.

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

[6] 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. Journal of the National Cancer Institute Monographs 2014; **2014**:187–197.

Selected References 2

[7] Kipourou DK, Perme MP, Rachet B, Belot A. Direct modeling of the crude probability of cancer death and the number of life years lost due to cancer without the need of cause of death: a pseudo-observation approach in the relative survival setting. Biostatistics Oxford England 2022;**23**:101–119.

[8] Lambert PC, Dickman PW, Nelson CP, Royston P. Estimating the crude probability of death due to cancer and other causes using relative survival models. Statistics in medicine 2010;**29**:885–895.

[9] Wong KF, Lambert PC, Mozumder SI, Broggio J, Rutherford MJ. Conditional crude probabilities of death for english cancer patients. British journal of cancer 2019;**121**:883–889.

[10] Rutherford MJ, Andersson TML, Myklebust TÅ, øller B, Lambert PC. Non-parametric estimation of reference adjusted, standardised probabilities of all-cause death and death due to cancer for population group comparisons. BMC medical research methodology 2022;**22**:2.

[11] Dasgupta P, Cramb SM, Kou K, Yu XQ, Baade PD. Quantifying the number of cancer deaths avoided due to improvements in cancer survival since the 1980s in the australian population, 1985-2014. Cancer epidemiology biomarkers prevention 2020;**29**:1825–1831.

[12] Jooste V, Bouvier AM, Bossard N, Uhry Z, Coureau G, Remontet L, et al.. Trends in probabilities of death owing to cancer and owing to other causes in patients with colon cancer. European journal of gastroenterology hepatology 2019;**31**:570–576.

Selected References 3

[13] Clèries R, Ameijide A, Buxó M, Martínez JM, Marcos-Gragera R, Vilardell ML, et al.. Long-term crude probabilities of death among breast cancer patients by age and stage: a population-based survival study in Northeastern Spain (Girona-Tarragona 1985-2004). Clinical translational oncology 2018;**20**:1252–1260.

[14] Orrason AW, Garmo H, Styrke J, Dickman PW, Stattin P. Comparison of relative survival and cause-specific survival in men with prostate cancer according to age and risk category: A nationwide, population-based study. American journal of epidemiology 2021;**190**:2053–2063.

[15] Wells M, Rutherford MJ, Lambert PC. Fair comparisons of cause-specific and relative survival by accounting for the systematic removal of patients from risk-sets. Cancer epidemiology 2023;**86**:102408.

[16] Makkar N, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. A comparison of relative survival and cause-specific survival methods to measure net survival in cancer populations. Cancer medicine 2018;**7**:4773–4780.

[17] Skyrud KD, Bray F, MÃ¸ller B. A comparison of relative and cause-specific survival by cancer site, age and time since diagnosis. International journal of cancer 2014;**135**:196–203.

[18] Withrow DR, Pole JD, Nishri ED, Tjepkema M, Marrett LD. Choice of relative or cause-specific approach to cancer survival analysis impacts estimates differentially by cancer type, population, and application: evidence from a canadian population-based cohort study. Population health metrics 2017;**15**:24.

Selected References 4

[19] Gamel JW, Vogel RL. Non-parametric comparison of relative versus cause-specific survival in surveillance, epidemiology and end results (seer) programme breast cancer patients. Statistical methods in medical research 2001;**10**:339–352.

[20] Sarfati D, Blakely T, Pearce N. Measuring cancer survival in populations: relative survival vs cancer-specific survival. International journal of epidemiology 2010;**39**:598–610.

[21] 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.

[22] Schaffar R, Rachet B, Belot A, Woods L. Cause-specific or relative survival setting to estimate population-based net survival from cancer? an empirical evaluation using women diagnosed with breast cancer in geneva between 1981 and 1991 and followed for 20 years after diagnosis. Cancer Epidemiol 2015;**39**:465–472.

[23] Schaffar R, Rapiti E, Rachet B, Woods L. Accuracy of cause of death data routinely recorded in a population-based cancer registry: impact on cause-specific survival and validation using the geneva cancer registry. BMC Cancer 2013;**13**:609.

[24] Schaffar R, Rachet B, Belot A, Woods LM. Estimation of net survival for cancer patients: Relative survival setting more robust to some assumption violations than cause-specific setting, a sensitivity analysis on empirical data. European journal of cancer Oxford England 1990 2017;**72**:78–83.

Selected References 5

[25] Pohar Perme M, Pavlic K. Nonparametric relative survival analysis with the R package relsurv. Journal of Statistical Software 2018;**87**:1–27.

[26] Pohar Perme M, Stare J, Estève J. On estimation in relative survival. Biometrics 2012;**68**:113–120.

[27] Corazziari I, Quinn M, Capocaccia R. Standard cancer patient population for age standardising survival ratios. Eur J Cancer 2004;**40**:2307–2316.

[28] Brenner H, Arndt V, Gefeller O, Hakulinen T. An alternative approach to age adjustment of cancer survival rates. Eur J Cancer 2004;**40**:2317–2322.

[29] Lambert PC, Andersson TML, Rutherford MJ, Myklebust TÅ, Møller B. Reference-adjusted and standardized all-cause and crude probabilities as an alternative to net survival in population-based cancer studies. International journal of epidemiology 2020;**49**:1614–1623.

[30] Rutherford MJ, Andersson TML, Myklebust Tg, MÃ¸ller B, Lambert PC. Non-parametric estimation of reference adjusted, standardised probabilities of all-cause death and death due to cancer for population group comparisons. BMC medical research methodology 2022;**22**:2.

Selected References 6

[31] Andersson TML, Rutherford MJ, MÃ¸ller B, Lambert PC, Myklebust Tg. Reference-adjusted loss in life expectancy for population-based cancer patient survival comparisons-with an application to colon cancer in sweden. Cancer epidemiology biomarkers prevention a publication of the American Association for Cancer Research cosponsored by the American Society of Preventive Oncology 2022;**31**:1720–1726.

[32] Myklebust TÅ, Aagnes B, Nilssen Y, Rutherford M, Lambert PC, Andersson TML, et al.. Improving communication of cancer survival statistics-feasibility of implementing model-based algorithms in routine publications. British journal of cancer 2023;**129**:819–828.

[33] De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, et al.. Cancer survival in europe 1999-2007 by country and age: results of eurocare–5-a population-based study. Lancet Oncol 2014;**15**:23–34.

[34] Syriopoulou E, Bower H, Andersson TL, Lambert P, Rutherford M. Estimating the impact of a cancer diagnosis on life expectancy by socio-economic group for a range of cancer types in England. British Journal of Cancer 2017; **117**:1419–1426.

[35] Howlader N, Mariotto AB, Woloshin S, Schwartz LM. Providing clinicians and patients with actual prognosis: cancer in the context of competing causes of death. J Natl Cancer Inst Monogr 2014;**2014**:255–264.

Selected References 7

[36] Forjaz de Lacerda G, Howlader N, Mariotto AB. Differences in cancer survival with relative versus cause-specific approaches: An update using more accurate life tables. Cancer epidemiology biomarkers prevention 2019; **28**:1544–1551.

[37] Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971-2011: a population-based study. Lancet 2014; **385**:1206–1218.

[38] Stewart SL, Harewood R, Matz M, Rim SH, Sabatino SA, Ward KC, Weir HK. Disparities in ovarian cancer survival in the united states (2001-2009): Findings from the concord-2 study. Cancer 2017;**123 Suppl 24**:5138–5159.

[39] Arnold M, Rutherford MJ, Bardot A, Ferlay J, Andersson TM, Myklebust TÅ, et al.. Progress in cancer survival, mortality, and incidence in seven high-income countries 1995-2014 (ICBP SURVMARK-2): a population-based study. Lancet Oncology 2019;**20**:1493–1505.

[40] Shack L, Bryant H, Lockwood G, Ellison LF. Conditional relative survival: a different perspective to measuring cancer outcomes. Cancer epidemiology 2013;**37**:446–448.

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

[42] Ellis L, Coleman MP, Rachet B. How many deaths would be avoidable if socioeconomic inequalities in cancer survival in England were eliminated? a national population-based study, 1996-2006. Eur J Cancer 2012;**48**:270–278.

Selected References 8

[43] Syriopoulou E, Rutherford M, Lambert P. Marginal measures and causal effects using the relative survival framework. International Journal of Epidemiology 2020;**49**:619–628.

[44] Rutherford MJ, Hinchliffe SR, Abel GA, Lyratzopoulos G, Lambert PC, Greenberg DC. How much of the deprivation gap in cancer survival can be explained by variation in stage at diagnosis: An example from breast cancer in the East of England. International Journal of Cancer 2013;**133**:2192–2200.

[45] Syriopoulou E, Rutherford MJ, Lambert PC. Understanding disparities in cancer prognosis: An extension of mediation analysis to the relative survival framework. Biometrical journal 2021;**63**:341–353.

[46] Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. Statistics in Medicine 2007;**26**:5486–5498.

[47] Royston P, Lambert PC. Flexible parametric survival analysis in Stata: Beyond the Cox model. Stata Press, 2011.

[48] Kipourou DK, Charvat H, Rachet B, Belot A. Estimation of the adjusted cause-specific cumulative probability using flexible regression models for the cause-specific hazards. Statistics in medicine 2019;**38**:3896–3910.

[49] Charvat H, Belot A. mexhaz: An r package for fitting flexible hazard-based regression models for overall and excess mortality with a random effect. Journal of Statistical Software 2021;**98**:1â“36.

[50] Jackson C. flexsurv: A platform for parametric survival modeling in r. Journal of Statistical Software 2016;**70**:1â“33.

www.encr.eu