Exploring inequalities in cancer prognosis across socioeconomic groups and how they arise

Elisavet Syriopoulou

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet

Email: elisavet.syriopoulou@ki.se

22nd November 2022

Outline

- About me
- Background on measuring cancer prognosis
  - Relative survival
  - Loss in life expectancy
- Differences in loss in life expectancy by socioeconomic group
- Investigating the determinants of observed differences
  - Lead time bias
  - Other factors (e.g. stage, treatment, comorbidity, etc)

About me

- BSc in Mathematics and MSc in Biostatistics
- 6-month research visit at the University of Copenhagen
- Moved to the UK and worked at the University of Leicester (October 2015)
- PhD in Biostatistics, University of Leicester (July 2020)
- In October 2020, I joined MEB as a postdoc working with Therese Andersson
- Current work focuses on the development and application of statistical methods for cancer registry data with a special focus on cancer disparities

Measuring cancer prognosis

- Relative survival: the proportion of patients who survive their disease by a specific time after diagnosis (net survival).
- Loss in life expectancy (LLE): the reduction in life expectancy following a cancer diagnosis.

Excess mortality rate and relative survival

It does not require cause of death information.

Excess mortality rate

$$\text{excess mortality} = \text{all-cause mortality} - \text{expected mortality}$$

It compares the all-cause mortality of the cancer population to the expected mortality of a comparable group in the general population.

Excess mortality and Relative survival - II

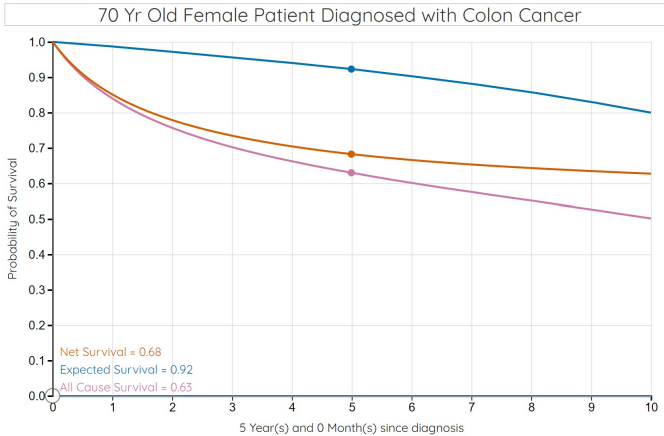
The survival analog of excess mortality is relative survival.

Relative survival

$$\text{relative survival} = \frac{\text{all-cause survival}}{\text{expected survival}}$$

The expected survival proportion is considered to be known and is usually obtained by available population lifetables.

Example - colon cancer, Sweden



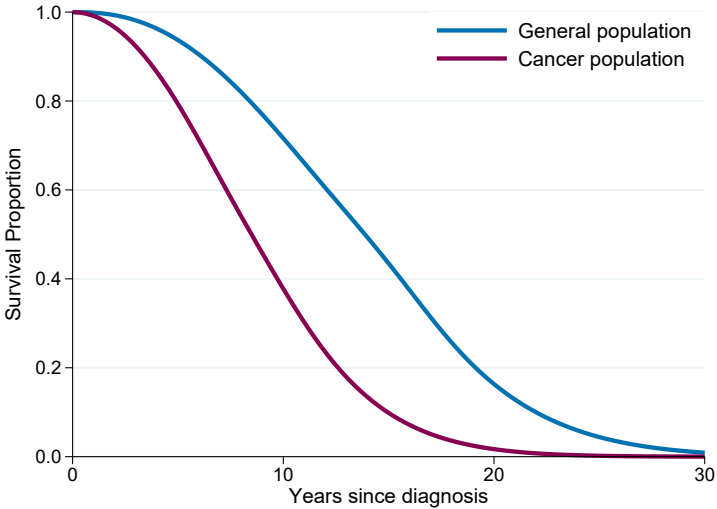
https://interpret.le.ac.uk

Loss in life expectancy

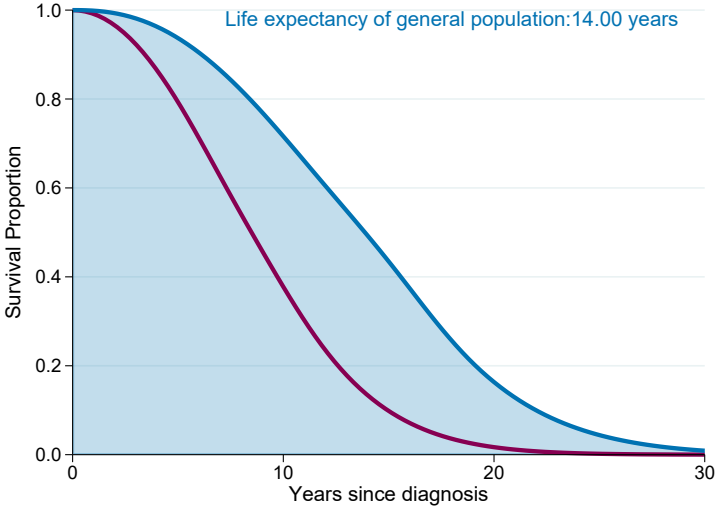
Loss in life expectancy (LLE): the difference between the life expectancy of the general population (assumed to be free from the cancer of interest) and the life expectancy of the cancer population (with similar characteristics).

$$\text{LLE}(\boldsymbol{Z} = \boldsymbol{z_i}) = \int_0^{t_{max}} \boldsymbol{S^*(t|Z_1 = z_{1i})}dt - \int_0^{t_{max}} \boldsymbol{S(t|Z = z_i)}dt$$

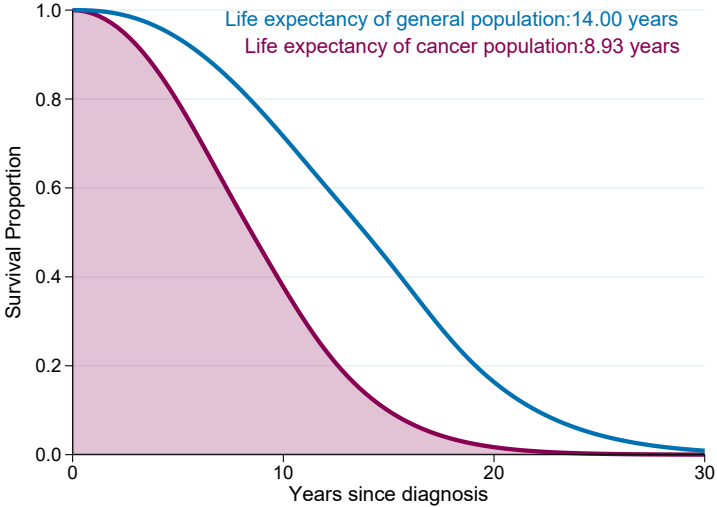
Loss in life expectancy



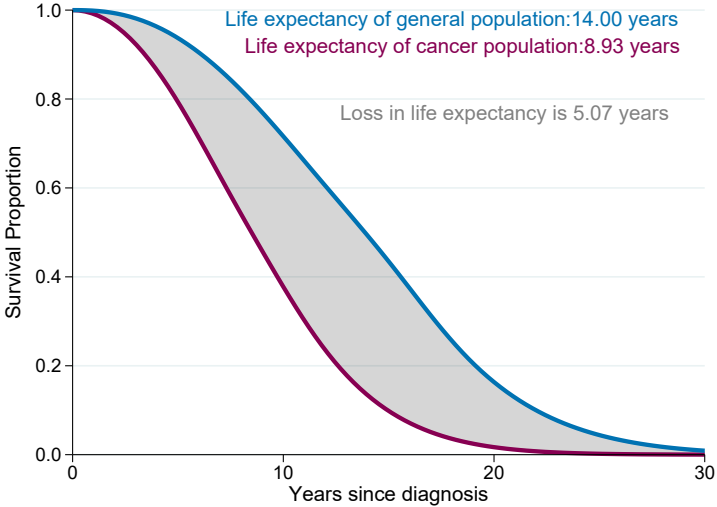
Loss in life expectancy



Loss in life expectancy



Loss in life expectancy



Estimation

$$\mathrm{LLE}(\boldsymbol{Z} = \boldsymbol{z_i}) = \int_0^{t_{max}} \boldsymbol{S^*(t|Z_1 = z_{1i})}dt - \int_0^{t_{max}} \boldsymbol{S(t|Z = z_i)}dt$$

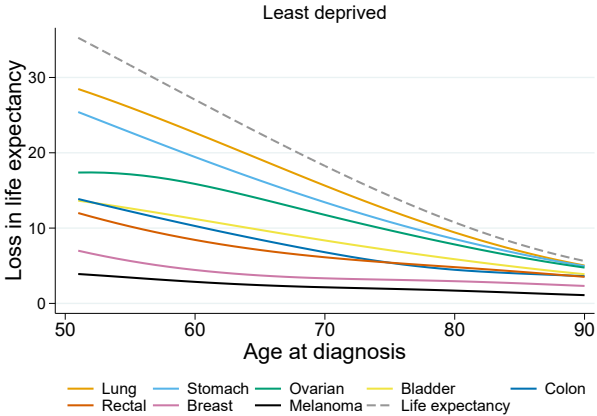
Due to limited follow-up, the survival curves should be extrapolated beyond available data.

Instead of the all-cause survival, it is easier to extrapolate the relative and expected survival:

$$\int_0^{t_{max}} \boldsymbol{S^*(t|Z_1 = z_{1i})}dt - \int_0^{t_{\max}} \boldsymbol{S^*(t|Z_1 = z_{1i})} \times \widehat{R}(t|\boldsymbol{Z_2 = z_{2i}})dt$$

Andersson T M-L, Dickman PW, Eloranta S, Lambe M, Lambert PC. Estimating the loss in expectation of life due to cancer using flexible parametric survival models. Stat Med 2013, 32:5286–5300.

Example - LLE by cancer type, England

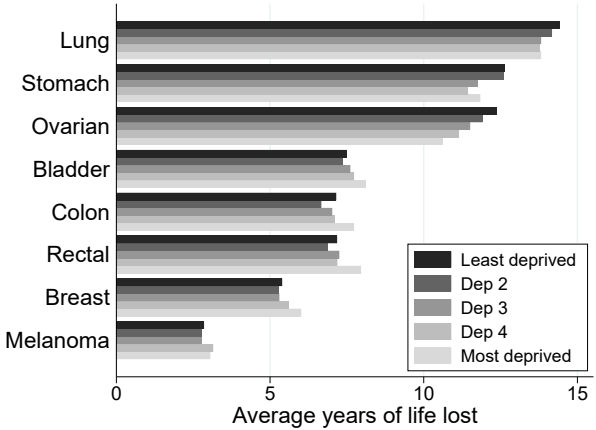


Syriopoulou E, Bower H, Andersson TM-L, Lambert PC, Rutherford MJ. Estimating the impact of a cancer diagnosis on life expectancy by socio-economic group for a range of cancer types in England. Br J Cancer 2017, 117:1419–1426, https://doi.org/10.1038/bjc.2017.300

Why use loss in life expectancy?

- LLE captures the cancer impact on the entire remaining lifespan.
- It can help us address useful questions:
  - Quantify the impact a cancer diagnosis has on a patient's life expectancy
  - Quantify disease burden in the society
    - "How many life-years are lost due to the cancer?"
    - "How many life-years are lost due to cancer by socioeconomic group?"

Example - LLE by SES groups, England



Syriopoulou E, Bower H, Andersson TM-L, Lambert PC, Rutherford MJ. Estimating the impact of a cancer diagnosis on life expectancy by socio-economic group for a range of cancer types in England. Br J Cancer 2017, 117:1419–1426, https://doi.org/10.1038/bjc.2017.300

Example - LLE by education groups, Sweden

Age (years)	Education group	Life years lost (95%
		Study population
45	Low	9.90 (8.46,11.33)
	Medium	9.69 (8.76,10.62)
	High	9.11 (8.10,10.13)
55	Low	5.42 (4.85,5.99)
	Medium	5.27 (4.78,5.76)
	High	5.03 (4.52,5.54)
65	Low	2.81 (2.53,3.08)
	Medium	2.69 (2.44,2.93)
	High	2.86 (2.54,3.18)
75	Low	2.28 (2.12,2.44)
	Medium	1.89 (1.74,2.05)
	High	2.11 (1.88,2.33)

Bower H, Andersson TM-L, Syriopoulou E, Rutherford MJ, Lambe M, Ahlgren J, Dickman PW, Lambert PC. Potential gain in life years for Swedish women with breast cancer if stage and survival differences between education groups could be eliminated - Three what-if scenarios. Breast 2019, 45:75–81, https://doi.org/10.1016/j.breast.2019.03.005

Understanding how inequalities arise is essential

- It can lead to a reduction of inequalities by targeting the most affected groups with relevant interventions.
- If survival differences across deprivation groups are largely driven by differences in stage at diagnosis, policies could be implemented to encourage earlier detection in the most deprived groups.
- If comorbidity accounts for some of the differences, this highlights the importance of focussing on other diseases to improving cancer outcomes as well.

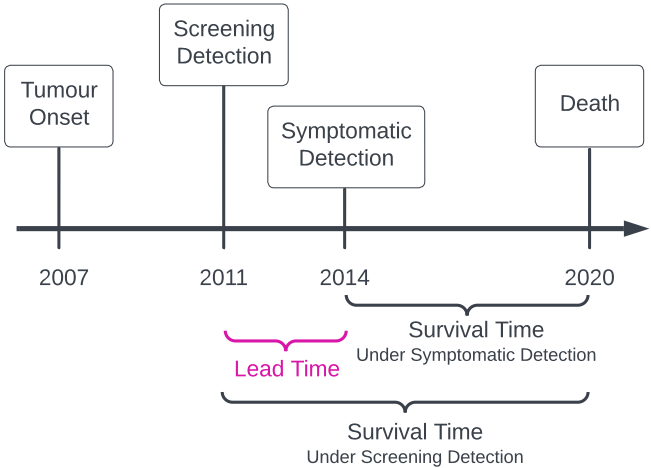
Why are there differences in LLE?

- Many factors have been suggested as potential drivers for the observed differences: stage at diagnosis, differential treatment, lifestyle, comorbidities, health-seeking behaviours, screening (and lead time bias), etc.

Why are there differences in LLE?

- Many factors have been suggested as potential drivers for the observed differences: stage at diagnosis, differential treatment, lifestyle, comorbidities, health-seeking behaviours, screening (and lead time bias), etc.
- The uptake of screening varies vary across socioeconomic groups, even in in countries where screening programmes are available on a national level.

Lead time


Tumour Onset
Screening Detection
Symptomatic Detection
Death
2007
2011
2014
2020
Survival Time
Under Symptomatic Detection
Lead Time
Survival Time
Under Screening Detection

Are LLE estimates affected by lead time bias?

- Screening can affect survival times both through real improvements in survival as well as an artificial increase (lead time bias) in survival times.

Are LLE estimates affected by lead time bias?

- Screening can affect survival times both through real improvements in survival as well as an artificial increase (lead time bias) in survival times.
- Partitioning the effect of screening into these two components is challenging as it would require knowledge of what would have happened in the absence of screening.

Assessing lead time bias

We use a simulation-based approach informed by Swedish cancer registry data which uses a natural history model developed in a Swedish setting.

Assessing lead time bias

We use a simulation-based approach informed by Swedish cancer registry data which uses a natural history model developed in a Swedish setting.

Syriopoulou *et al. Breast Cancer Research* (2022) 24:15
https://doi.org/10.1186/s13058-022-01505-3

Breast Cancer Research

RESEARCH ARTICLE Open Access

Assessing lead time bias due to mammography screening on estimates of loss in life expectancy

Elisavet Syriopoulou*, Alessandro Gasparini, Keith Humphreys and Therese M.-L. Andersson

Simulation

Simulated data with age at tumour onset based on incidence rates in Sweden from 1973 (the year before the introduction of mammography screening).

Simulation

Simulated data with age at tumour onset based on incidence rates in Sweden from 1973 (the year before the introduction of mammography screening).

For individuals with onset of breast cancer:

- Tumour growth (and time to symptomatic detection) was simulated from a natural history model developed in a Swedish setting

Abrahamsson L and Humphreys K. A statistical model of breast cancer tumour growth with estimation of screening sensitivity as a function of mammographic density. Statistical Methods in Medical Research 2016, 25:1620–1637.

Simulation

Simulated data with age at tumour onset based on incidence rates in Sweden from 1973 (the year before the introduction of mammography screening).

For individuals with onset of breast cancer:

- Tumour growth (and time to symptomatic detection) was simulated from a natural history model developed in a Swedish setting
- Time to death was simulated as the minimum between:
  - time to death due to cancer
  - time to death due to other causes

Abrahamsson L and Humphreys K. A statistical model of breast cancer tumour growth with estimation of screening sensitivity as a function of mammographic density. Statistical Methods in Medical Research 2016, 25:1620–1637.

Screening scenarios

We also imposed a mammography screening programme with individuals invited to screening every second year between ages 40–74.

Screening scenarios

We also imposed a mammography screening programme with individuals invited to screening every second year between ages 40–74.

- Screening sensitivity
  - Low
  - Moderate
  - High

Screening scenarios

We also imposed a mammography screening programme with individuals invited to screening every second year between ages 40–74.

- Screening sensitivity
  - Low
  - Moderate
  - High
- Screening attendance
  - Perfect (everyone attends all screening visits)
  - Imperfect
    - 80% of the individuals attend each scheduled screen visit with a probability of 90%
    - 20% of the individuals attend each scheduled screen visit with a probability of 15%

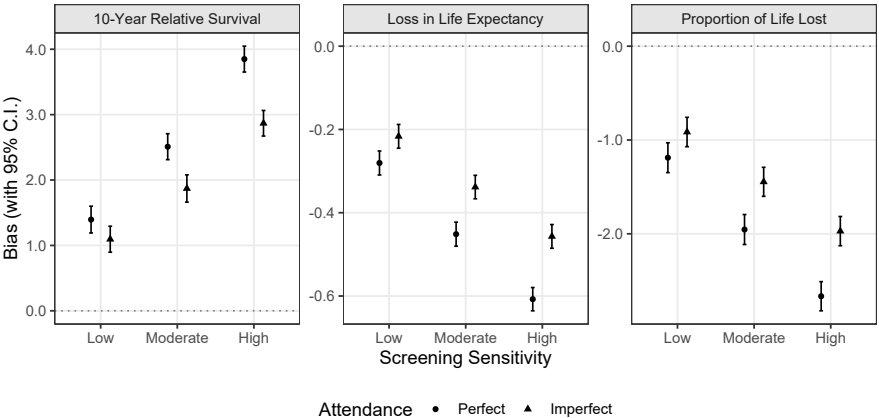
Lead time bias

We compare

- LLE estimates calculated in the absence of screening (*actual* value)
- LLE estimates when a screening programme is imposed

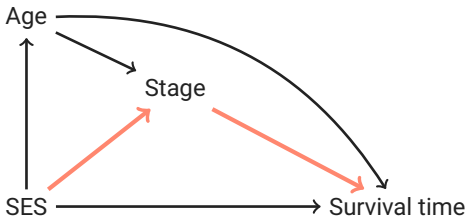
Here the only difference is that screen detected tumours result in an earlier diagnosis. The actual survival time was not changed for screen detected cases!

Bias


10-Year Relative Survival
Loss in Life Expectancy
Proportion of Life Lost
Bias (with 95% C.I.)
4.0
3.0
2.0
1.0
0.0
0.0
-0.2
-0.4
-0.6
0.0
-1.0
-2.0
Low
Moderate
High
Screening Sensitivity
Attendance
Perfect
Imperfect

Investigating other factors

Could stage at diagnosis partly explain the survival differences between the least and most deprived groups?



This is a mediation analysis question!

Mediation analysis

Mediation analysis methods allow to explore the role of a third variable (mediator) on an observed association between an exposure and an outcome of interest.

Mediation analysis

Mediation analysis methods allow to explore the role of a third variable (mediator) on an observed association between an exposure and an outcome of interest.

Complex mechanisms contribute towards cancer disparities:

- Cancer-related factors,
- Other cause factors.

Mediation analysis

Mediation analysis methods allow to explore the role of a third variable (mediator) on an observed association between an exposure and an outcome of interest.

Complex mechanisms contribute towards cancer disparities:

- Cancer-related factors,
- Other cause factors.

During my PhD, I adapted a formal causal framework to the settings of cancer registry-based epidemiology, extending mediation analysis methods to the relative survival framework.

Mediation analysis

Mediation analysis methods allow to explore the role of a third variable (mediator) on an observed association between an exposure and an outcome of interest.

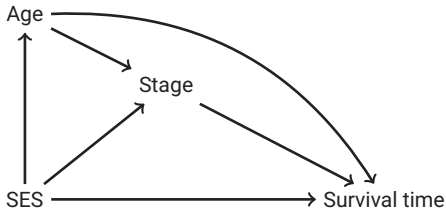
Complex mechanisms contribute towards cancer disparities:

- Cancer-related factors,
- Other cause factors.

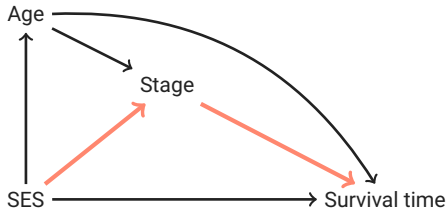
During my PhD, I adapted a formal causal framework to the settings of cancer registry-based epidemiology, extending mediation analysis methods to the relative survival framework.

Main idea: using the relative survival framework allows to isolate cancer-related factors.

Partitioning the total survival difference



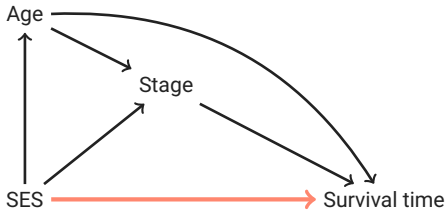
Partitioning the total survival difference



Natural indirect effect: quantifies how much of the observed difference is due to stage differences in the two groups

$$\text{NIE}(t) = R^{1,M^1}(t) - R^{1,M^0}(t) = E[R(t|X=1,\mathbf{Z_2},M^1)] - E[R(t|X=1,\mathbf{Z_2},M^0)]$$

Partitioning the total survival difference



Natural indirect effect: quantifies how much of the observed difference is due to stage differences in the two groups

$$\text{NIE}(t) = R^{1,M^1}(t) - R^{1,M^0}(t) = E[R(t|X=1,\mathbf{Z_2},M^1)] - E[R(t|X=1,\mathbf{Z_2},M^0)]$$

Natural direct effect: quantifies the differences in relative survival that are not due to stage differences

$$\text{NDE}(t) = R^{1,M^0}(t) - R^{0,M^0}(t) = E[R(t|X=1,\mathbf{Z_2},M^0)] - E[R(t|X=0,\mathbf{Z_2},M^0)]$$

ESTIMATION

$$\widehat{\text{NDE}}(t) = \frac{1}{N}\sum_{i=1}^{N}\sum_{m}\hat{R}(t|X=1,\mathbf{Z_2}=\mathbf{z_{2i}},M=m)\hat{P}(M=m|X=0,\mathbf{Z_2}=\mathbf{z_{2i}})$$
$$-\frac{1}{N}\sum_{i=1}^{N}\sum_{m}\hat{R}(t|X=0,\mathbf{Z_2}=\mathbf{z_{2i}},M=m)\hat{P}(M=m|X=0,\mathbf{Z_2}=\mathbf{z_{2i}})$$

$$\widehat{\text{NIE}}(t) = \frac{1}{N}\sum_{i=1}^{N}\sum_{m}\hat{R}(t|X=1,\mathbf{Z_2}=\mathbf{z_{2i}},M=m)\hat{P}(M=m|X=1,\mathbf{Z_2}=\mathbf{z_{2i}})$$
$$-\frac{1}{N}\sum_{i=1}^{N}\sum_{m}\hat{R}(t|X=1,\mathbf{Z_2}=\mathbf{z_{2i}},M=m)\hat{P}(M=m|X=0,\mathbf{Z_2}=\mathbf{z_{2i}})$$

Received: 25 November 2019 | Revised: 7 August 2020 | Accepted: 21 September 2020

DOI: 10.1002/bimj.201900355

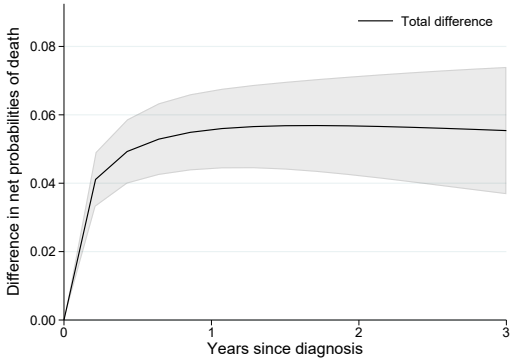
Biometrical Journal

RESEARCH PAPER

Understanding disparities in cancer prognosis: An extension of mediation analysis to the relative survival framework

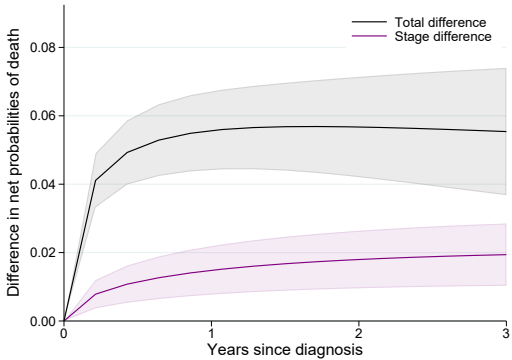
Elisavet Syriopoulou[1] | Mark J. Rutherford[1] | Paul C. Lambert[1,2]

Example - colon cancer, England



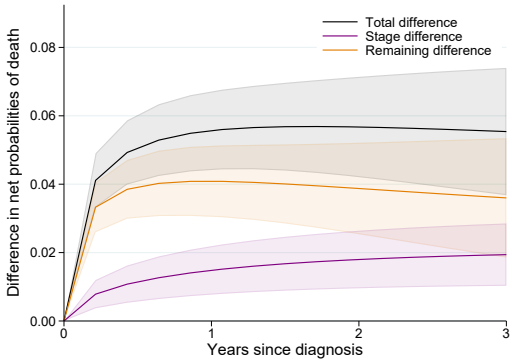
At 3-years after diagnosis, the total difference between the least and most deprived is 5.5 percentage points.

Example - colon cancer, England



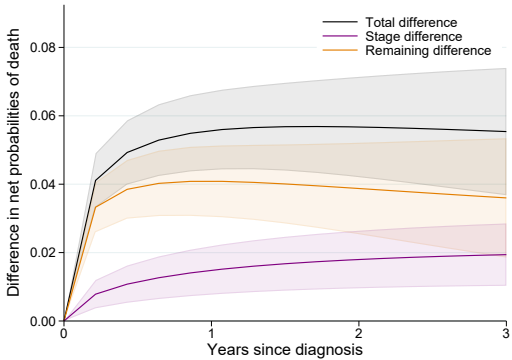
From the total difference, 1.9 percentage points are due to stage differences.

Example - colon cancer, England



From the total difference, the remaining 3.6 percentage points are due to other factors.

Example - colon cancer, England



At 3-years after diagnosis, stage explains 35% (=1.9/5.5) of the total differences.

Hypothetical interventions

- We saw that 35% of the survival differences between deprivations groups are due to stage differences.

Hypothetical interventions

- We saw that 35% of the survival differences between deprivations groups are due to stage differences.
- How much would survival improve if we could shift the stage distribution of the most deprived to that of the least deprived?

Hypothetical interventions

- We saw that 35% of the survival differences between deprivations groups are due to stage differences.
- How much would survival improve if we could shift the stage distribution of the most deprived to that of the least deprived?
- We can think about it as assessing the impact of a potential intervention aimed at eliminating stage-related differences.

Hypothetical interventions

- We saw that 35% of the survival differences between deprivations groups are due to stage differences.
- How much would survival improve if we could shift the stage distribution of the most deprived to that of the least deprived?
- We can think about it as assessing the impact of a potential intervention aimed at eliminating stage-related differences.
- This is what we do with mediation analysis, assessing the impact of interventions.

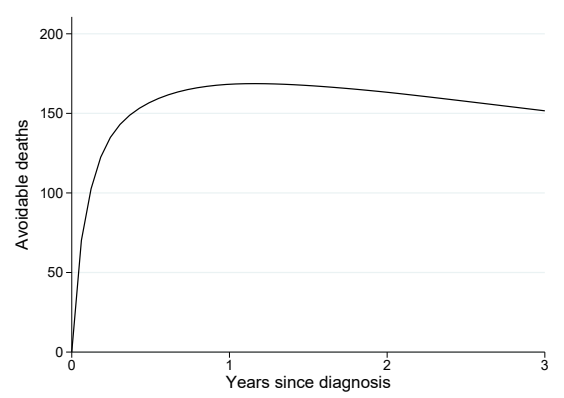
Hypothetical interventions

- We saw that 35% of the survival differences between deprivations groups are due to stage differences.
- How much would survival improve if we could shift the stage distribution of the most deprived to that of the least deprived?
- We can think about it as assessing the impact of a potential intervention aimed at eliminating stage-related differences.
- This is what we do with mediation analysis, assessing the impact of interventions.
- The impact of such interventions can also be quantified in terms of avoidable (postponable) deaths.

Avoidable deaths

How many “avoidable deaths” would there be if the most deprived had the same stage distribution as the least deprived?

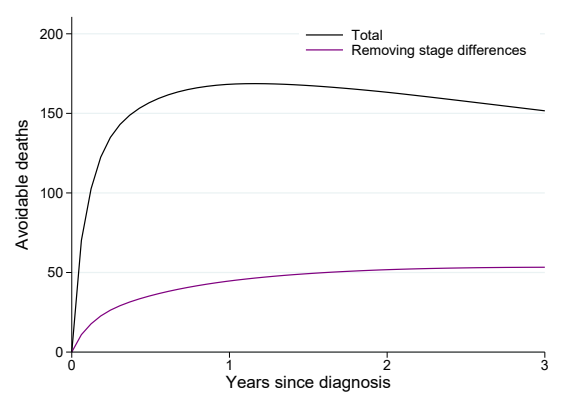
Avoidable deaths - Colon cancer, England



* Out of 3228 patients from the most deprived group diagnosed in 2013

At 3-years after diagnosis, there are 151 total avoidable deaths:

Avoidable deaths - Colon cancer, England

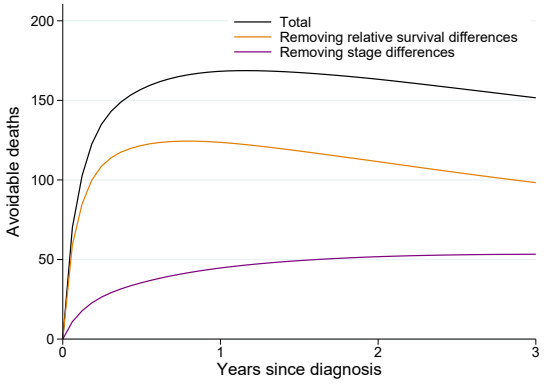


* Out of 3228 patients from the most deprived group diagnosed in 2013

At 3-years after diagnosis, there are 151 total avoidable deaths:

53 are by eliminating stage differences &

Avoidable deaths - Colon cancer, England



* Out of 3228 patients from the most deprived group diagnosed in 2013

At 3-years after diagnosis, there are 151 total avoidable deaths:

53 are by eliminating stage differences &

98 are by eliminating remaining differences

Next steps: Swedish data and FORTE

- We will explore differences in cancer prognosis (e.g. loss in life expectancy) for various cancer types in Sweden using data from cancer quality registers.
- Recently we obtained data on colorectal cancer (CRC BaSe) and for this we will collaborate with Caroline Nordenvall and Erik Osterman.
- For melanoma, we are currently recruiting a PhD student and we will work closely with Hanna Eriksson.
- We will need to further extend the methods to accommodate the complexities of the available data.

How to quantify SES?

Highest achieved educational level as an indicator.

- However, the school system constantly changes: need to account for temporal changes.
- Sensitivity analyses by restricting analysis to birth cohorts that have undergone schooling within the same system.

How to quantify SES?

Highest achieved educational level as an indicator.

- However, the school system constantly changes: need to account for temporal changes.
- Sensitivity analyses by restricting analysis to birth cohorts that have undergone schooling within the same system.

Income as an indicator.

- However, a substantial part of our study population consists of older age groups, in which the income information might be problematic.

How to quantify SES?

Highest achieved educational level as an indicator.

- However, the school system constantly changes: need to account for temporal changes.
- Sensitivity analyses by restricting analysis to birth cohorts that have undergone schooling within the same system.

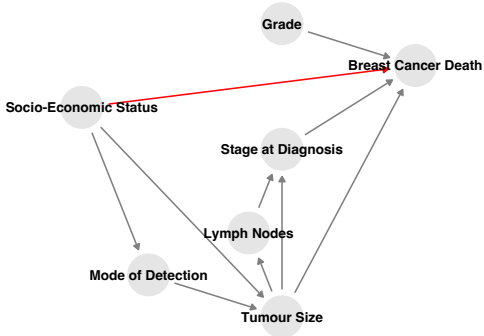
Income as an indicator.

- However, a substantial part of our study population consists of older age groups, in which the income information might be problematic.

We will work closely with Alexander Miething (Sociologist and Researcher, Stockholm University).

How to categorise stage at diagnosis?

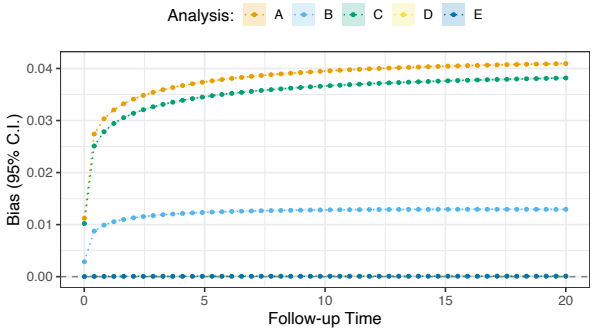
- Stage at diagnosis (I to IV) may not be able to fully capture the heterogeneity between patients - it is likely too crude.
- Using more detailed categories (T-N-M cross-product) may be better at capturing heterogeneity.



EXAMPLE - BREAST CANCER

We conducted a microsimulation experiment (with A Gasparini, K Humphreys).

Generated data with no (direct) survival difference between high & low SES groups. Mediation analysis using stage (I-III) yield differences up to 4%.



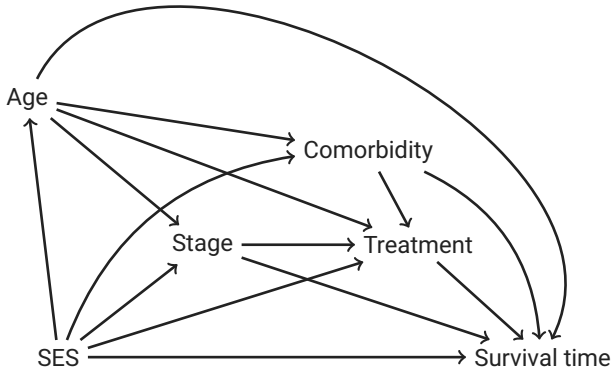
Constructing population lifetables by SES

- We will need population lifetables for the expected mortality rates startified by age, sex, calendar period and SES.
- However, the available lifetables don't have information on SES.
- When data are not available on a population level, information from a control population can be used to adjust expected rates.
- Bower et al. suggested an approach using a Poisson generalized linear model or a flexible parametric survival model.

Bower H, Andersson TM-L,Crowther MJ, Dickman PW, Lambe M, Lambert PC. Adjusting Expected Mortality Rates Using Information From a Control Population: An Example Using Socioeconomic Status, American Journal of Epidemiology 2018, 187(4):828–836.

It will get trickier!

There will be multiple mediators (stage, treatment, comorbidity, etc).



Limited interventions

- What if an intervention cannot fully eliminate the differences in the stage distribution?
- For instance, a specific intervention might increase early diagnosis only in certain age groups (e.g. only those within screening ages).
- What impact would these limited interventions have in terms of life-years gained?

Other things to consider

- Missing data and how to deal with these properly (multiple imputation approaches);
- Computation time: combining MI with bootstrap for obtaining standard errors;
- Communication of results.

Conclusions

- There are differences in the prognosis of cancer patients.

Conclusions

- There are differences in the prognosis of cancer patients.
- Understanding mechanisms driving disparities is important.

Conclusions

- There are differences in the prognosis of cancer patients.
- Understanding mechanisms driving disparities is important.
- Understanding mechanisms driving disparities is difficult!

Conclusions

- There are differences in the prognosis of cancer patients.
- Understanding mechanisms driving disparities is important.
- Understanding mechanisms driving disparities is difficult!
- Causal mediation analysis can be a valuable tool for exploring such settings.