



**Karolinska  
Institutet**

# An extension of mediation analysis to the relative survival framework

Elisavet Syriopoulou

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet

Email: [elisavet.syriopoulou@ki.se](mailto:elisavet.syriopoulou@ki.se)

22nd February 2024

# Outline

- Background
  - Excess mortality and relative survival
  - Socioeconomic differences in cancer survival
- Mediation analysis and measures of interest
- Assumption for identification
- Estimation in Stata
- Comments and thoughts

# Background

- I work mainly with cancer registry data where interest is on cancer-specific survival.
- A common issue is the presence of competing events.
- Focus is often on estimating a measure called *net* survival
  - i.e. survival in a hypothetical world where the only possible cause of death is the event of interest e.g. death due to cancer
  - i.e. the survival that would be observed if it was possible to eliminate all competing causes of death
- This hypothetical construct allows comparisons of cancer survival between populations without distortions from differential background mortality
- There are two approaches for estimating net survival:
  - cause-specific survival
  - relative survival

# Background

- I work mainly with cancer registry data where interest is on cancer-specific survival.
- A common issue is the presence of competing events.
- Focus is often on estimating a measure called *net survival*
  - i.e. survival in a hypothetical world where the only possible cause of death is the event of interest e.g. death due to cancer
  - i.e. the survival that would be observed if it was possible to eliminate all competing causes of death
- This hypothetical construct allows comparisons of cancer survival between populations without distortions from differential background mortality
- There are two approaches for estimating net survival:
  - cause-specific survival
  - **relative survival**

## 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

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

$$R(t) = \frac{S(t)}{S^*(t)} \quad \text{and} \quad S(t) = S^*(t)R(t)$$

The expected survival proportions (or expected mortality rates) are considered to be known and are usually obtained by available population lifetables, stratified by factors such as age, sex, calendar time etc.

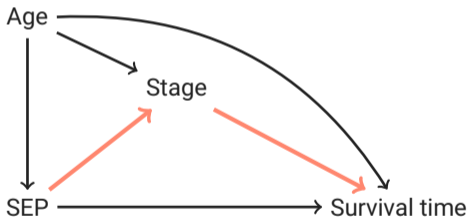
## Motivation - Example, England

Survival after a cancer diagnosis varies considerably across population groups e.g socioeconomic groups.

Deprivation Group	5-year RS	Mean Years w/o Cancer	Mean Years with Cancer	Prop (%)
<hr/>				
Age at diagnosis: 70				
Least deprived	63.02	16.08	9.81	38.99
2	61.26	15.12	9.11	39.78
3	60.03	14.49	8.59	40.73
4	56.90	13.56	7.78	42.60
Most deprived	54.74	12.66	6.94	45.15

## 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 allows to explore the role of a mediator on an observed association between an exposure - outcome of interest.

## Mediation analysis

Mediation analysis allows to explore the role of a mediator on an observed association between an exposure - outcome of interest.

However, we still need to deal with the complex mechanisms that contribute towards cancer disparities:

- Cancer-related factors
- Other cause factors

## Mediation analysis

Mediation analysis allows to explore the role of a mediator on an observed association between an exposure - outcome of interest.

However, we still need to deal with the complex mechanisms that contribute towards cancer disparities:

- Cancer-related factors
- Other cause factors

We have adapted a formal causal framework to the settings of cancer registry-based epidemiology, extending mediation analysis methods to the relative survival framework<sup>1</sup>.

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

<sup>1</sup> 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.

## Notation

- $X$  denotes the exposure of interest (here SEP), with  $x = 1$  if exposed and  $x = 0$  if unexposed
- $M$  denotes the mediator of interest (here stage)
- $Z$  denotes the confounding variables (here age)
- $R(t)$  denotes relative survival at time  $t$

## Notation - potential outcomes

The framework used for formulating the effects of interest is that of potential outcomes: *the outcomes that would be observed if we intervene on  $X$  and  $M$  to set them on specific values.*

## Notation - potential outcomes

The framework used for formulating the effects of interest is that of potential outcomes: *the outcomes that would be observed if we intervene on  $X$  and  $M$  to set them on specific values.*

- $R^x(t)$  is the value that  $R(t)$  would have if we intervened on  $X$  and set it (possibly counter to fact) to the value  $x$
- $M^x$  is the value that  $M$  would take if we intervened on  $X$  and set it to  $x$
- $R^{x, M^{x^*}}(t)$  is the the value that  $R(t)$  would take if we intervened on  $X$  and set it to  $x$  and simultaneously intervened on  $M$  and set it to  $M^{x^*}$ , where  $x$  and  $x^*$  are not necessarily the same.

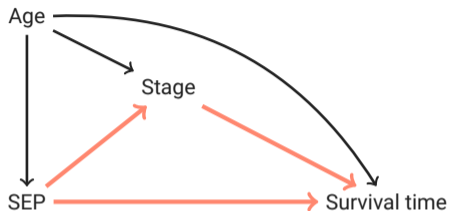
## Notation - potential outcomes

The framework used for formulating the effects of interest is that of potential outcomes: *the outcomes that would be observed if we intervene on  $X$  and  $M$  to set them on specific values.*

- $R^x(t)$  is the value that  $R(t)$  would have if we intervened on  $X$  and set it (possibly counter to fact) to the value  $x$
- $M^x$  is the value that  $M$  would take if we intervened on  $X$  and set it to  $x$
- $R^{x,M^{x^*}}(t)$  is the the value that  $R(t)$  would take if we intervened on  $X$  and set it to  $x$  and simultaneously intervened on  $M$  and set it to  $M^{x^*}$ , where  $x$  and  $x^*$  are not necessarily the same.

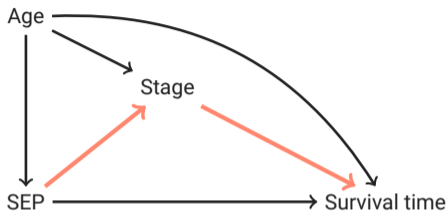
We will use those to define contrasts of marginal effects of the potential outcomes.

## Partitioning the total survival difference





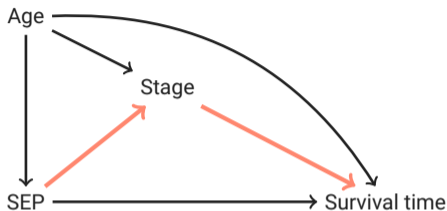
## Partitioning the total survival difference



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

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

## Partitioning the total survival difference



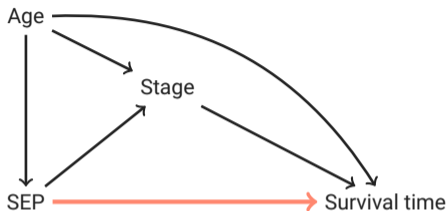
Natural indirect effect: quantifies how much of the observed difference is due to stage differences in the two groups

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

This is a comparison of two hypothetical quantities:

- In the first,  $M$  is set to  $M^1$  and in the second  $M$  is set to  $M^0$ , for everyone. In both,  $X$  is set to 1.
- $X$  influence  $R(t)$  only through its influence on  $M$ .
- Thus corresponds to an indirect effect through  $M$ .

## Partitioning the total survival difference



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

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

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

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

## Comment

The definitions involve the following term:

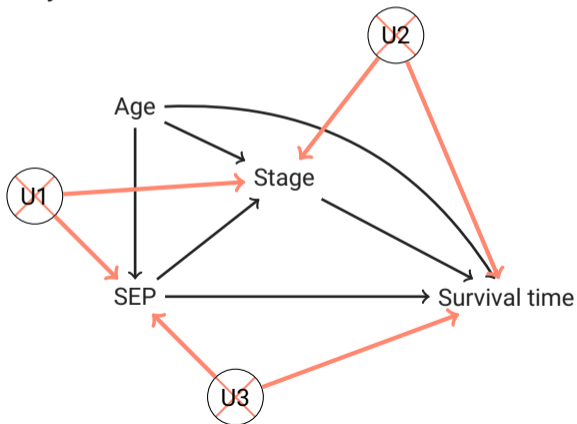
$$R(t|X = 1, \mathbf{Z}, M^0)$$

i.e., the relative survival if setting the exposure to the level of the exposed and the mediator to the mediator value if unexposed!

An additional model is required for the mediator.

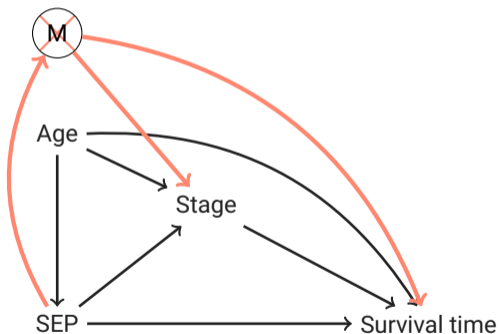
## Assumptions for identification

To link the hypothetical quantities with the observed data, we need to assume no interference, consistency and conditional exchangeability.



## Assumptions for identification

We also need to make the cross-world independence assumption that implies that there are no intermediate confounders  $M$ .



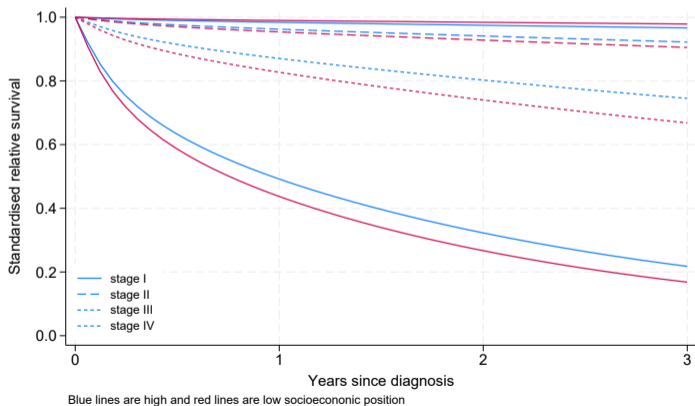
## Estimation - regression standardisation

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

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

## Example - simulated colon cancer data

I will use an example of simulated colon cancer data that is available with the methodological paper.





## Identification algorithm

1. Fit a relative survival model including  $X, M, \mathbf{Z}$ .
2. Fit model for the mediator including  $X, \mathbf{Z}$ .
3. For each individual in the study population obtain estimates for  $\hat{P}(M = m|X = x, \mathbf{Z} = \mathbf{z}_i)$ , at each  $X = x$ .
4. Obtain estimates of standardised  $\hat{R}(t|X = x, \mathbf{Z} = \mathbf{z}_i, M = m)$  at  $X = x$ , using the predictions of Step 3 as weights. Form contrasts to obtain the  $\widehat{NDE}_{RS}$  and  $\widehat{NIE}_{RS}$ .

## Identification algorithm

1. Fit a relative survival model including  $X, M, \mathbf{Z}$ .
2. Fit model for the mediator including  $X, \mathbf{Z}$ .
3. For each individual in the study population obtain estimates for  $\hat{P}(M = m|X = x, \mathbf{Z} = \mathbf{z}_i)$ , at each  $X = x$ .
4. Obtain estimates of standardised  $\hat{R}(t|X = x, \mathbf{Z} = \mathbf{z}_i, M = m)$  at  $X = x$ , using the predictions of Step 3 as weights. Form contrasts to obtain the  $\widehat{NDE}_{RS}$  and  $\widehat{NIE}_{RS}$ .
5. Repeat from Step 3,  $k$  times, while performing parametric bootstrap for the parameter estimates for both models.
6. Calculate 95% confidence intervals either by taking the 2.5% and 97.5% percentiles of the estimates across the bootstrapped samples or by using their standard deviation.

# Estimation in Stata

The following Stata user-written commands will be needed

- `stpm3`
- `gensplines`
- `standsurv`

# Estimation in Stata

## 1. Fit a relative survival model including $X$ , $M$ , $Z$ .

```
// Merge in expected mortality rates using the population lifetable
gen _age = min(int(agediag+_t),99)
gen _year = min(floor(yeardiag + _t),2016)
sort sex _year _age dep
merge m:1 sex _year _age dep using popmort.dta, ///
      keep(match master) keepusing(rate)
sort id

// Declare survival data
stset t, failure(dead==1) id(id) exit(time 3)
//Fit a flexible parametric survival model
stpm3 i.dep i.sex @ns(agediag, df(3)) i.stage i.dep##i.stage, ///
      tvc(i.sex @ns(agediag, df(3))) dftvc(2) ///
      scale(lncumhazard) df(4) bhazard(rate)

estimates store surv
```

# Estimation in Stata

## 2. Fit a separate model for the mediator including $X$ , $Z$ .

```
// Generate splines for age
gensplines ageddiag, df(3) type(ns) gen(ns_age)
local ageknots 'r(knots)'
```

```
// Fit a multinomial regression model for the low SEP
// (exposed/most deprived)
mlogit stage ns_age* i.sex if dep==1
estimates store ph1
```

```
// Fit a multinomial regression model for the high SEP
// (unexposed/least deprived)
mlogit stage ns_age** i.sex if dep==0
estimates store ph0
```

## Estimation in Stata

3. For each individual in the study population obtain estimates for  $\hat{P}(M = m|X = x, \mathbf{Z} = \mathbf{z}_i)$ , at each  $X = x$ .

```
// For low SEP
estimates restore ph1
gen tmpdep = dep
replace dep = 1
predict p11 p12 p13 p14
replace dep = tmpdep
drop tmpdep

// For high SEP
estimates restore ph0
gen tmpdep = dep
replace dep = 0
predict p01 p02 p03 p04
replace dep = tmpdep
drop tmpdep
```

## Estimation in Stata

4. Obtain estimates of standardised  $\hat{R}(t|X = x, \mathbf{Z} = \mathbf{z}_i, M = m)$  at  $X = x$ , using the estimates of Step 3 as weights.

$$TCE_{RS} = E [R(t|X = 1, \mathbf{Z}, M^1)] - E [R(t|X = 0, \mathbf{Z}, M^0)]$$

```
range tt 0 3 50 //Generate time variable for predictions

estimates restore surv

standsurv, timevar(tt) failure frame(med, replace) ///
    at1(dep 1 stage 1, atindweights(p11)) ///
    at2(dep 1 stage 2, atindweights(p12)) ///
    at3(dep 1 stage 3, atindweights(p13)) ///
    at4(dep 1 stage 4, atindweights(p14)) ///
    at5(dep 0 stage 1, atindweights(p01)) ///
    at6(dep 0 stage 2, atindweights(p02)) ///
    at7(dep 0 stage 3, atindweights(p03)) ///
    at8(dep 0 stage 4, atindweights(p04)) ///
lincom(1 1 1 1 -1 -1 -1 -1) lincomvar(tce_PE)
```

Use the option `failure` to predict net probabilities of death (1-relative survival).

## Estimation in stata

$$NDE_{RS} = E [R(t|X = 1, \mathbf{Z}, \mathbf{M}^0)] - E [R(t|X = 0, \mathbf{Z}, \mathbf{M}^0)]$$



## Estimation in stata

$$NDE_{RS} = E [R(t|\mathbf{X} = \mathbf{1}, \mathbf{Z}, M^0)] - E [R(t|\mathbf{X} = \mathbf{0}, \mathbf{Z}, M^0)]$$

```
standsurv, timevar(tt) failure frame(med, merge)   ///
    at1(dep 1 stage 1, atindweights(p01))         ///
    at2(dep 1 stage 2, atindweights(p02))         ///
    at3(dep 1 stage 3, atindweights(p03))         ///
    at4(dep 1 stage 4, atindweights(p04))         ///
    at5(dep 0 stage 1, atindweights(p01))         ///
    at6(dep 0 stage 2, atindweights(p02))         ///
    at7(dep 0 stage 3, atindweights(p03))         ///
    at8(dep 0 stage 4, atindweights(p04))         ///
    lincom(1 1 1 1 -1 -1 -1 -1) lincomvar(tde_PE)
```

## Estimation in Stata

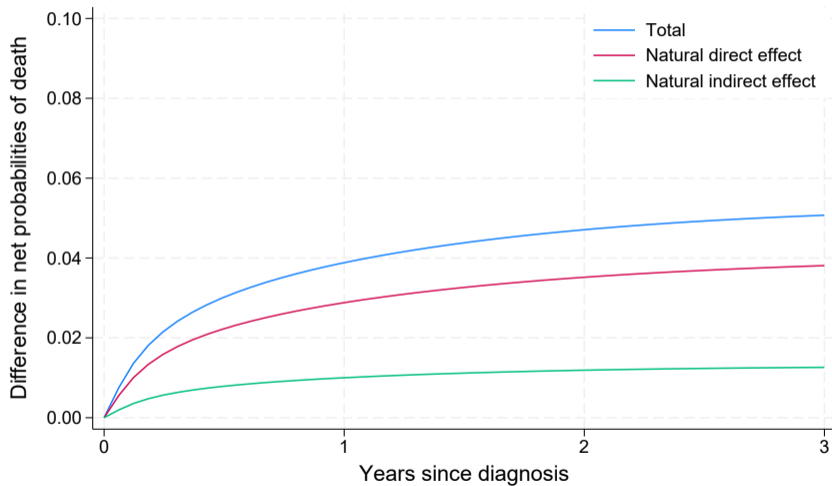
$$NIE_{RS} = E [R(t|\mathbf{X} = \mathbf{1}, \mathbf{Z}, M^1)] - E [R(t|\mathbf{X} = \mathbf{1}, \mathbf{Z}, M^0)]$$

## Estimation in Stata

$$NIE_{RS} = E [R(t|X = 1, \mathbf{Z}, \mathbf{M}^1)] - E [R(t|X = 1, \mathbf{Z}, \mathbf{M}^0)]$$

```
standsurv, timevar(tt) failure frame(med, merge) ///
    at1(dep 1 stage 1, atindweights(p11))      ///
    at2(dep 1 stage 2, atindweights(p12))      ///
    at3(dep 1 stage 3, atindweights(p13))      ///
    at4(dep 1 stage 4, atindweights(p14))      ///
    at5(dep 1 stage 1, atindweights(p01))      ///
    at6(dep 1 stage 2, atindweights(p02))      ///
    at7(dep 1 stage 3, atindweights(p03))      ///
    at8(dep 1 stage 4, atindweights(p04))      ///
    lincom(1 1 1 1 -1 -1 -1 -1) lincomvar(tie_PE)
```

## Example - plot



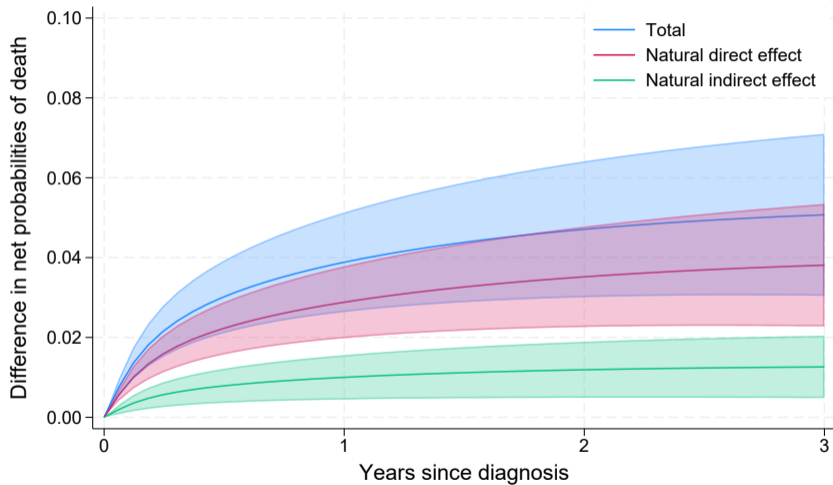
## Confidence intervals

- We need to account for the uncertainty on both the probabilities estimated in Step 3 and the survival functions of Step 4.
- Bootstrap-based standard errors will be obtained.
- By performing parametric bootstrap, the parameters are drawn repeatedly from a multivariate normal distribution and for each draw we obtain both estimates and the variance covariance matrix that are finally combined.

There is an example available on this on GitHub:

<https://github.com/syriop-elisa/mediation-example-stpm3>

## Plot - with CIs



## Standardising among subsets

- We can also define contrasts within specific subsets such as the NDE and NIE among the exposed (e.g. among the most deprived)

$$NIE_{exposed} = E \left[ R(t|X = 1, \mathbf{Z}^{*X=1}, M^1) \right] - E \left[ R(t|X = 1, \mathbf{Z}^{*X=1}, M^0) \right]$$

- For the estimation, we restrict standardisation to a specific subset of the population.

## Moving to a real-world setting

- So far, we only talked about a net setting that requires elimination of competing events and looked at
  - differences in relative survival
- Differences can also be quantified in a real-world setting where no elimination of competing events is required
  - difference in all-cause survival
  - avoidable deaths
- To do so, we need to incorporate the expected mortality rates,  $S^*(t)$ , in the contrast of interest.
  - There are many different ways to do it.



## Direct & indirect effects - all cause setting

- Use the observed distribution of the exposure for  $S^*(t)$ :

$$NIE_{AC} = E \left[ S^*(t|X, Z^*) R(t|X = 1, Z^{*X=1}, M^1) \right] \\ - E \left[ S^*(t|X, Z^*) R(t|X = 1, Z^{*X=1}, M^0) \right]$$

*Differences in all-cause survival can only be due to the cancer of interest (not due to other-cause differences).*

- The difference in all-cause survival between the two SEP groups if we could intervene and shift the stage distribution of the lowest SEP group to that of the highest SEP group, while keeping their background mortality unchanged?

## Direct & indirect effects - all cause setting

- Use the observed distribution of the exposure for  $S^*(t)$ :

$$\begin{aligned} NIE_{AC} = E & \left[ S^*(t|X, Z^*) R(t|X = 1, Z^{*X=1}, M^1) \right] \\ & - E \left[ S^*(t|X, Z^*) R(t|X = 1, Z^{*X=1}, M^0) \right] \end{aligned}$$

*Differences in all-cause survival can only be due to the cancer of interest (not due to other-cause differences).*

- The difference in all-cause survival between the two SEP groups if we could intervene and shift the stage distribution of the lowest SEP group to that of the highest SEP group, while keeping their background mortality unchanged?

A conceptually similar measure has recently been proposed for the standard survival setting, so called separable effects:

```

standsurv if dep==1 , timevar(tt) failure frame(med, merge)   ///
    at1(dep 1 stage 1, atindweights(p11))                ///
    at2(dep 1 stage 2, atindweights(p12))                ///
    at3(dep 1 stage 3, atindweights(p13))                ///
    at4(dep 1 stage 4, atindweights(p14))                ///
    at5(dep 1 stage 1, atindweights(p01))                ///
    at6(dep 1 stage 2, atindweights(p02))                ///
    at7(dep 1 stage 3, atindweights(p03))                ///
    at8(dep 1 stage 4, atindweights(p04))                ///
lincom(1 1 1 1 -1 -1 -1 -1) lincomvar(tie_PE_ac) ///
expssurv(using(popmort.dta)                               ///
    datediag(datediag)                                   ///
    ageddiag(agediag)                                   ///
    pmrate(rate)                                         ///
    pmage(_age)                                          ///
    pmyear(_year)                                        ///
    pmmayear(2016)                                       ///
    pmother(dep sex)                                    ///
    at1(.)                                               ///
    at2(.)                                               ///
    at3(.)                                               ///
    at4(.)                                               ///
    at5(.)                                               ///
    at6(.)                                               ///
    at7(.)                                               ///
    at8(.)                                               ///
)

```

## Avoidable deaths under hypothetical interventions

- So far, we have looked at survival differences (or probabilities of death) under certain interventions, e.g. shifting the mediator distribution of one group to that of another.
- The impact of such interventions can also be quantified as the avoidable deaths.
- For instance, how many deaths would be postponed for the lowest SEP group under an intervention that aims to shift the distribution of stage at diagnosis for the lowest SEP to that of the highest SEP group, while keeping other cause mortality rates unchanged/constant?

## Avoidable deaths

- The predicted number of deaths for the exposed  $X = 1$  in a typical calendar year with  $N^*$  diagnoses:

$$N^* \times \left(1 - E \left[ S^*(t|X, \mathbf{Z}^{*X=1}) R(t|X = 1, \mathbf{Z}^{X=1}, \mathbf{M}^1) \right]\right)$$

- The expected number of deaths under the intervention of shifting the mediator distribution of the exposed to the one of the unexposed (setting  $M$  to  $M^0$ ):

$$N^* \times \left(1 - E \left[ S^*(t|X, \mathbf{Z}^{*X=1}) R(t|X = 1, \mathbf{Z}^{X=1}, \mathbf{M}^0) \right]\right)$$

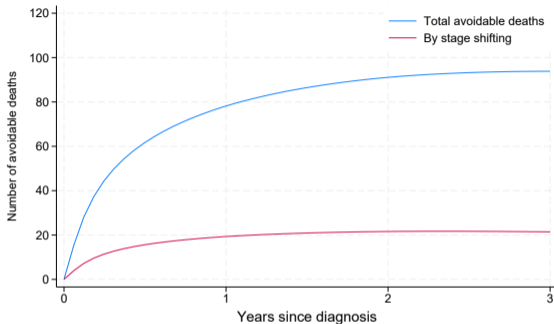
- The avoidable deaths is given by their difference.
- We keep the expected survival of the lowest SEP group constant and we assess an intervention that aims to shift the stage distribution of the lowest SEP to that of the high SEP with no impact on other cause mortality rates.*

```

stands surv if dep==1, per(2170) timevar(tt) failure frame(med, merge) ///
    at1(dep 1 stage 1, atindweights(p11))          ///
    at2(dep 1 stage 2, atindweights(p12))          ///
    at3(dep 1 stage 3, atindweights(p13))          ///
    at4(dep 1 stage 4, atindweights(p14))          ///
    at5(dep 1 stage 1, atindweights(p01))          ///
    at6(dep 1 stage 2, atindweights(p02))          ///
    at7(dep 1 stage 3, atindweights(p03))          ///
    at8(dep 1 stage 4, atindweights(p04))          ///
lincom(1 1 1 1 -1 -1 -1 -1) lincomvar(ADs_PE) ///
expsurv(using(popmort.dta) ///
    datediag(datediag)                            ///
    ageddiag(agediag)                             ///
    pmrate(rate)                                  ///
    pmage(_age)                                   ///
    pmyear(_year)                                 ///
    pmmayear(2016)                               ///
    pmother(dep sex)                             ///
    at1(.)                                        ///
    at2(.)                                        ///
    at3(.)                                        ///
    at4(.)                                        ///
    at5(.)                                        ///
    at6(.)                                        ///
    at7(.)                                        ///
    at8(.)                                        ///
)

```

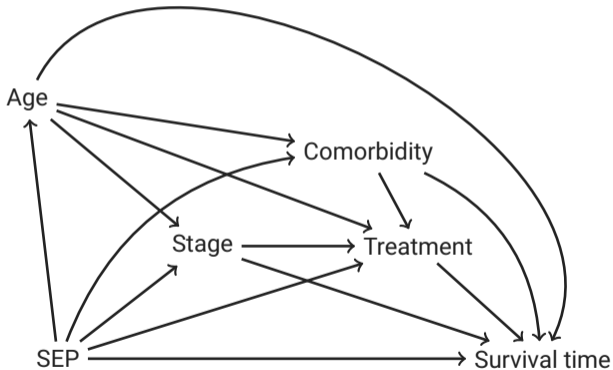
# Avoidable deaths for colon cancer



- For total ADs, we shifted the relative survival and stage at diagnosis distribution of the lowest SEP to that of the highest SEP.
- For both scenarios, we kept the expected survival of the lowest SEP unchanged.
- 3 years after diagnosis there would be approx. 94 avoidable deaths in total, out of 2170 patients from the lowest SEP diagnosed in 2013 the most recent year in our cohort study.
- Partitioning that further, we found that approx. 22 deaths of the total deaths would be from eliminating stage differences and the remaining 72 would be from removing the remaining relative survival differences.

## It will get trickier!

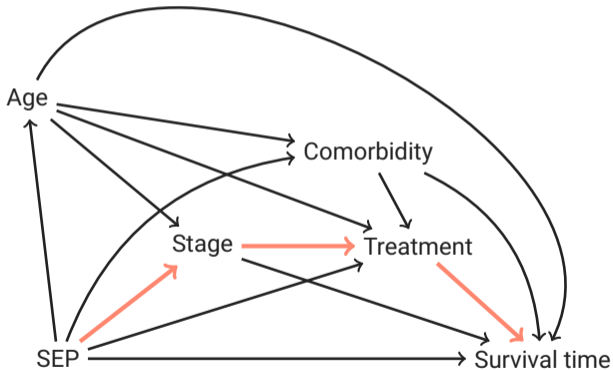
There will be multiple mediators (stage, treatment, comorbidity, etc).





## It will get trickier!

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



## Settings with multiple mediators

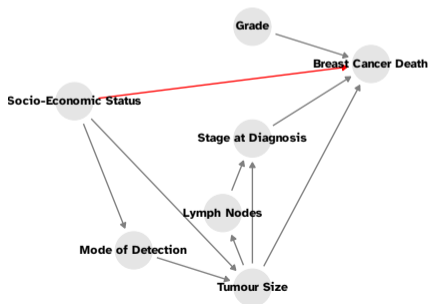
- Randomised interventional analogues of the NDE and the NIE have been proposed and these require no cross-world assumption<sup>1,2</sup>.
- The proposed effect decomposition is relatively easy to perform via a (Monte-Carlo based) regression approach. It delivers effects mediated via each of the mediators separately, but also via the mediators' dependence.
- We are currently working to extend the above approach into the relative survival framework.

<sup>1</sup> VanderWeele TJ, Vansteelandt S, Robins, James M. Effect Decomposition in the Presence of an Exposure-Induced Mediator-Outcome Confounder. *Epidemiology* 2014.

<sup>2</sup> Vansteelandt S, Daniel RM. Interventional Effects for Mediation Analysis with Multiple Mediators. *Epidemiology* 2017.

# Imperfectly Defined Mediators: stage categorisation

- Stage at diagnosis (I to IV) may not be able to fully capture the heterogeneity between patients - it is likely too crude.
- Preliminary results from a microsimulation of breast cancer, showed that using more detailed categories (T-N-M cross-product) may be better at capturing heterogeneity.



# 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 using relative survival can be a valuable tool for exploring such settings.