Survival differences between socioeconomic groups of colon cancer patients: exploring the role of stage at diagnosis using mediation analysis

Elisavet Syriopoulou[1], Mark Rutherford[2], Paul Lambert[1,2]

[1]Department of Medical Epidemiology and Biostatistics,
  Karolinska Institutet, Sweden
[2]Biostatistics Research Group, Department of Health Sciences,
  University of Leicester, UK

E-mail: elisavet.syriopoulou@ki.se
November 2020

MOTIVATION

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 (%)
Age at diagnosis: 70				
Least deprived	63.6	18.3	11.4	37.7
2	62.3	17.4	10.7	38.5
3	60.5	16.8	10.0	40.3
4	58.8	16.1	9.4	41.6
Most deprived	54.0	15.4	8.2	46.5

[1]Understanding the impact of socioeconomic differences in colorectal cancer survival: potential gain in life-years. *Brit J Cancer* 2019; 20:1052–1058

MOTIVATION

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 (%)
Age at diagnosis: 70				
Least deprived	63.6	18.3	11.4	37.7
2	62.3	17.4	10.7	38.5
3	60.5	16.8	10.0	40.3
4	58.8	16.1	9.4	41.6
Most deprived	54.0	15.4	8.2	46.5

[1]Understanding the impact of socioeconomic differences in colorectal cancer survival: potential gain in life-years. *Brit J Cancer* 2019; 20:1052–1058

Motivation

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 (%)
Age at diagnosis: 70				
Least deprived	63.6	18.3	11.4	37.7
2	62.3	17.4	10.7	38.5
3	60.5	16.8	10.0	40.3
4	58.8	16.1	9.4	41.6
Most deprived	54.0	15.4	8.2	46.5

[1]Understanding the impact of socioeconomic differences in colorectal cancer survival: potential gain in life-years. *Brit J Cancer* 2019; 20:1052–1058

MOTIVATION

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 (%)
Age at diagnosis: 70				
Least deprived	63.6	18.3	11.4	37.7
2	62.3	17.4	10.7	38.5
3	60.5	16.8	10.0	40.3
4	58.8	16.1	9.4	41.6
Most deprived	54.0	15.4	8.2	46.5

[1]Understanding the impact of socioeconomic differences in colorectal cancer survival: potential gain in life-years. *Brit J Cancer* 2019; 20:1052–1058

Motivation

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 (%)
Age at diagnosis: 70				
Least deprived	63.6	18.3	11.4	37.7
2	62.3	17.4	10.7	38.5
3	60.5	16.8	10.0	40.3
4	58.8	16.1	9.4	41.6
Most deprived	54.0	15.4	8.2	46.5

[1]Understanding the impact of socioeconomic differences in colorectal cancer survival: potential gain in life-years. *Brit J Cancer* 2019; 20:1052–1058

Understanding variation in relative survival

Is there a third variable that can partly explain these differences?

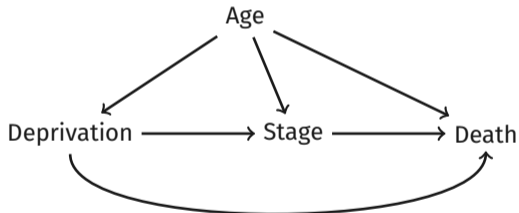# Understanding variation in relative survival

Is there a third variable that can partly explain these differences?

Data on 23639 patients diagnosed between 2011-2013 (41.46% in the most deprived group).

Stage at diagnosis	Least deprived	Most deprived
I	1338 (9.70)	912 (9.27)
II	2644 (19.16)	1950 (19.81)
III	2435 (17.65)	1716 (17.43)
IV	2585 (18.74)	2050 (20.83)
Missing	4794 (34.75)	3215 (32.66)

UNDERSTANDING VARIATION IN RELATIVE SURVIVAL

Is there a third variable that can partly explain these differences?

Data on 23639 patients diagnosed between 2011-2013 (41.46% in the most deprived group).

Stage at diagnosis	Least deprived	Most deprived
I	1338 (9.70)	912 (9.27)
II	2644 (19.16)	1950 (19.81)
III	2435 (17.65)	1716 (17.43)
IV	2585 (18.74)	2050 (20.83)
Missing	4794 (34.75)	3215 (32.66)

Could stage at diagnosis partly explain the survival differences?

THE ROLE OF STAGE AT DIAGNOSIS

Could stage at diagnosis partly explain the survival differences?
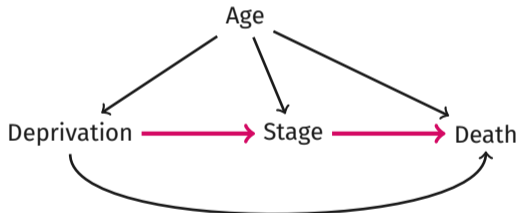
Methodological considerations

Mediation analysis methods allow the exploration of the role of a third variable that may explain part of the differences.

Methodological considerations

Mediation analysis methods allow the exploration of the role of a third variable that may explain part of the differences.

However, complex mechanisms contribute towards disparities. All-cause survival differences are the result of both:

- cancer-related factors
- other cause factors

Methodological considerations

Mediation analysis methods allow the exploration of the role of a third variable that may explain part of the differences.

However, complex mechanisms contribute towards disparities. All-cause survival differences are the result of both:

- cancer-related factors
- other cause factors

The relative survival framework allows isolation of cancer-related differences , the determinants of which might be easier to identify.

Methodological considerations

Mediation analysis methods allow the exploration of the role of a third variable that may explain part of the differences.

However, complex mechanisms contribute towards disparities. All-cause survival differences are the result of both:

- cancer-related factors
- other cause factors

The relative survival framework allows isolation of cancer-related differences , the determinants of which might be easier to identify.

Mediation analysis into the relative survival framework can be a valuable tool!

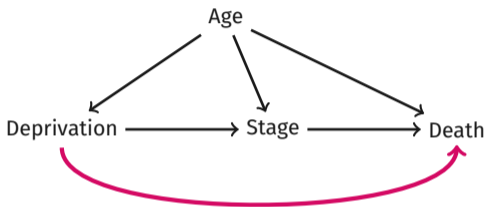# Partitioning the total relative survival difference
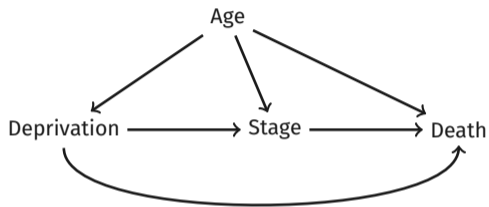
PARTITIONING THE TOTAL RELATIVE SURVIVAL DIFFERENCE
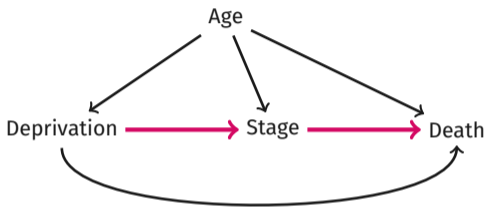
Natural indirect effect: quantifies how much of the observed difference is due to stage differences in the two groups

PARTITIONING THE TOTAL RELATIVE SURVIVAL DIFFERENCE

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

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

Statistical models

- Fitted a flexible parametric relative survival model that allows more flexibility on capturing the underlying baseline excess hazard.

Statistical models

- Fitted a flexible parametric relative survival model that allows more flexibility on capturing the underlying baseline excess hazard.
- The model included sex, deprivation, age and stage at diagnosis.

Statistical models

- Fitted a flexible parametric relative survival model that allows more flexibility on capturing the underlying baseline excess hazard.
- The model included sex, deprivation, age and stage at diagnosis.
- Allowed for time-dependent effects for deprivation, age and stage as well as an interaction between stage and deprivation.

Statistical models

- Fitted a flexible parametric relative survival model that allows more flexibility on capturing the underlying baseline excess hazard.
- The model included sex, deprivation, age and stage at diagnosis.
- Allowed for time-dependent effects for deprivation, age and stage as well as an interaction between stage and deprivation.
- A separate model was fitted for stage at diagnosis: a multinomial regression model including age, deprivation and sex.

Statistical models

- Fitted a flexible parametric relative survival model that allows more flexibility on capturing the underlying baseline excess hazard.
- The model included sex, deprivation, age and stage at diagnosis.
- Allowed for time-dependent effects for deprivation, age and stage as well as an interaction between stage and deprivation.
- A separate model was fitted for stage at diagnosis: a multinomial regression model including age, deprivation and sex.
- Missing information on stage was imputed using multiple imputations with 35 imputed dataset.

Statistical models

- Fitted a flexible parametric relative survival model that allows more flexibility on capturing the underlying baseline excess hazard.
- The model included sex, deprivation, age and stage at diagnosis.
- Allowed for time-dependent effects for deprivation, age and stage as well as an interaction between stage and deprivation.
- A separate model was fitted for stage at diagnosis: a multinomial regression model including age, deprivation and sex.
- Missing information on stage was imputed using multiple imputations with 35 imputed dataset.
- To obtain 95% confidence intervals parametric bootstraps of 200 iterations were performed for each imputed dataset.
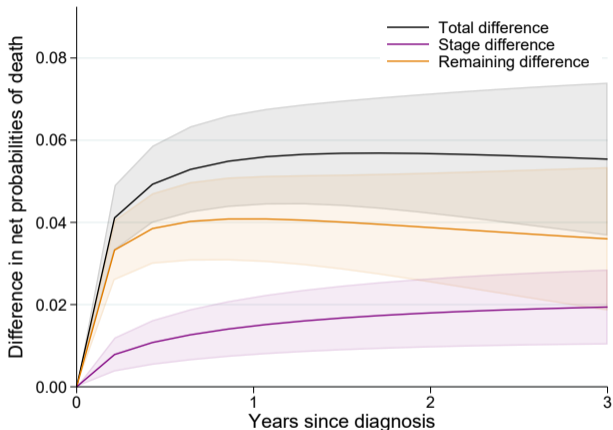
PARTITIONING THE TOTAL DIFFERENCE
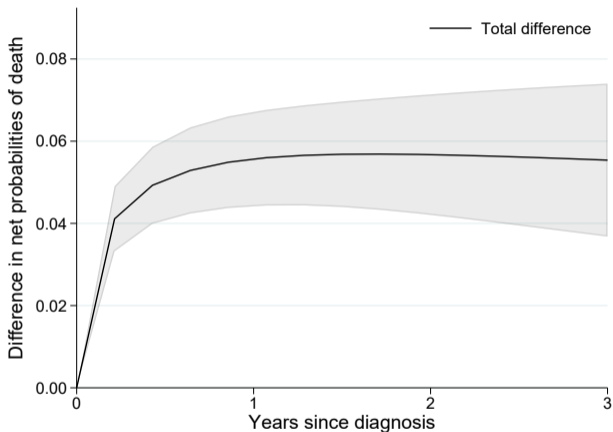
At 3-years after diagnosis, the total difference is **5.5** percentage points.
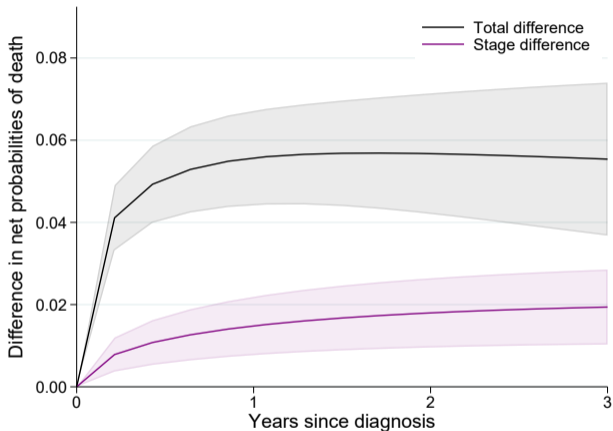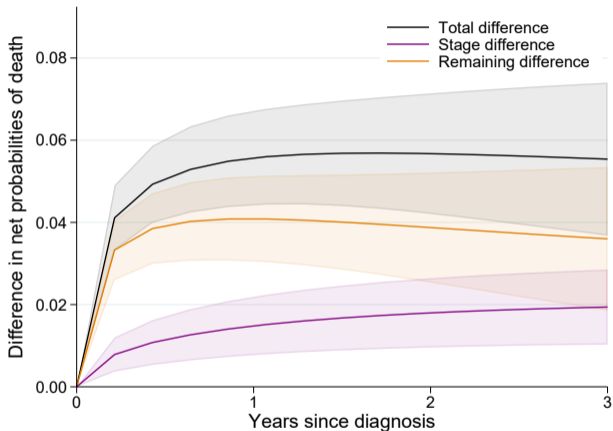
From the total difference, **1.9** percentage points are due to stage differences.

PARTITIONING THE TOTAL DIFFERENCE

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

PARTITIONING THE TOTAL DIFFERENCE

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

Moving to a real-world setting

- Net setting refers to hypothetical world where colon cancer is the only possible cause of death

Moving to a real-world setting

- Net setting refers to hypothetical world where colon cancer is the only possible cause of death
- Differences can also be quantified in a real-word setting where other causes of death are present:
  - difference in all-cause probabilities of death
  - avoidable deaths

Moving to a real-world setting

- Net setting refers to hypothetical world where colon cancer is the only possible cause of death
- Differences can also be quantified in a real-word setting where other causes of death are present:
    - difference in all-cause probabilities of death
    - avoidable deaths

How many "avoidable deaths" would there be if the most deprived had the same stage distribution as the least deprived?
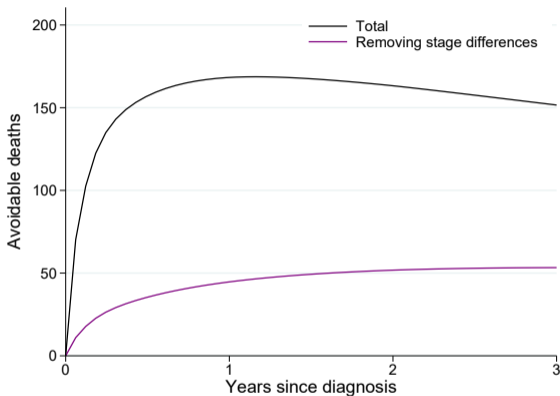
AVOIDABLE DEATHS WHEN REMOVING STAGE DIFFERENCES
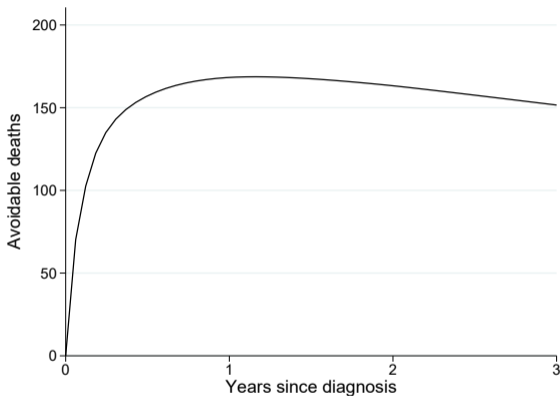
*Out of 3228 patients from the most deprived group diagnosed in 2013

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

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

At 3-years after diagnosis, there are **151** total avoidable deaths:
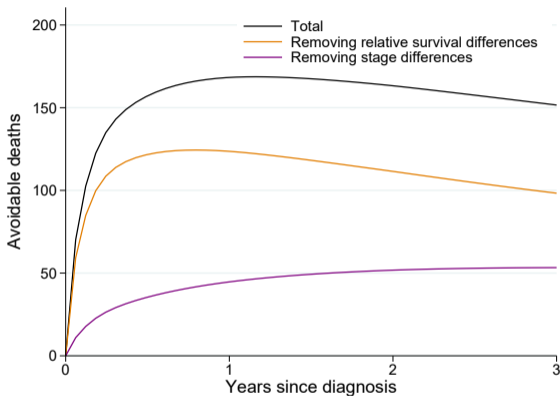   **53** are by eliminating stage differences &

AVOIDABLE DEATHS WHEN REMOVING STAGE DIFFERENCES

*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

Conclusions

- The proportion of differences that is explained by stage is increasing with time and an intervention aiming to eliminate stage differences would result in a substantial number of avoidable deaths.

Conclusions

- The proportion of differences that is explained by stage is increasing with time and an intervention aiming to eliminate stage differences would result in a substantial number of avoidable deaths.
- Understanding which factors drive differences is essential:

Conclusions

- The proportion of differences that is explained by stage is increasing with time and an intervention aiming to eliminate stage differences would result in a substantial number of avoidable deaths.
- Understanding which factors drive differences is essential:
  - It can lead to a reduction of inequalities by targeting the most affected groups with relevant interventions.

Conclusions

- The proportion of differences that is explained by stage is increasing with time and an intervention aiming to eliminate stage differences would result in a substantial number of avoidable deaths.
- Understanding which factors drive differences is essential:
  - It can lead to a reduction of inequalities by targeting the most affected groups with relevant interventions.
  - For example, if survival differences across deprivation groups are largely driven by differences in stage at diagnosis, then policies could be implemented to encourage earlier detection in the most deprived groups.

Conclusions

- The proportion of differences that is explained by stage is increasing with time and an intervention aiming to eliminate stage differences would result in a substantial number of avoidable deaths.
- Understanding which factors drive differences is essential:
  - It can lead to a reduction of inequalities by targeting the most affected groups with relevant interventions.
  - For example, if survival differences across deprivation groups are largely driven by differences in stage at diagnosis, then policies could be implemented to encourage earlier detection in the most deprived groups.
- Mediation analysis into the relative survival framework provides an opportunity to answer these complex questions.

Selected References

Pohar Perme M, Stare J & Estève J
On Estimation in Relative Survival.
Biometrics,, 68, 113-120, 2012.

Royston P & Lambert P C
Flexible parametric survival analysis in Stata: Beyond the Cox model.
Stata Press, 2011.

Syriopoulou E, Rutherford M J & Lambert, P C
Marginal measures and causal effects using the relative survival framework.
International Journal of Epidemiology, 49, 619–628, 2020.

Syriopoulou E, Rutherford M J & Lambert, P C
Understanding disparities in cancer prognosis: An extension of mediation analysis to the relative survival framework.
Biometrical Journal, in press, 2020.