How to estimate cancer survival

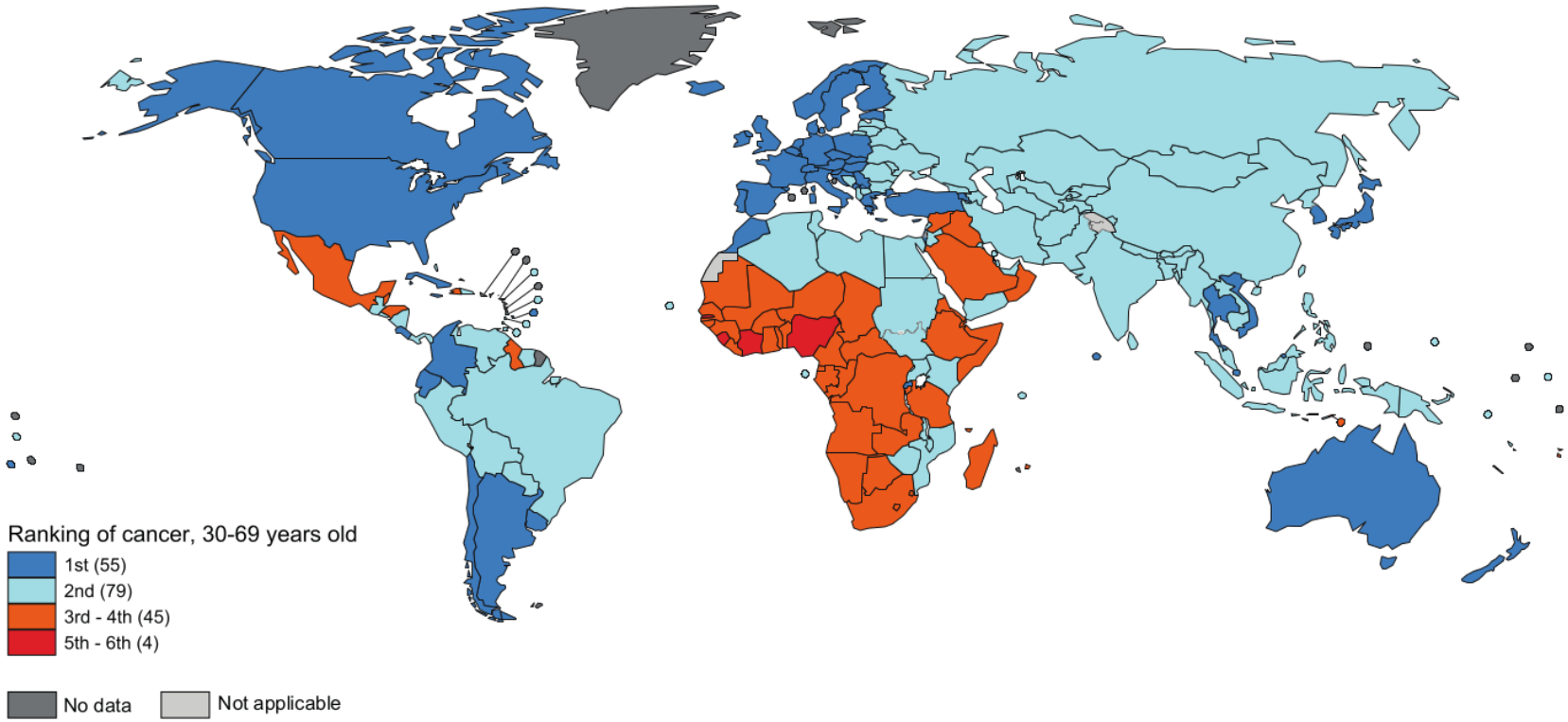
(an overview)

Javier González-Delgado

Institut de Mathématiques de Toulouse, LAAS-CNRS

March 24, 2022

The global cancer burden



Global map of cancer as a leading cause of premature death (i.e. at ages 30-69 years)

Figure: Wild C.P., Weiderpass E., Stewart B.W., editors (2020). *World Cancer Report: Cancer Research for Cancer Prevention.* Lyon, France: International Agency for Research on Cancer.

The *unequal* cancer burden

Inequalities in cancer

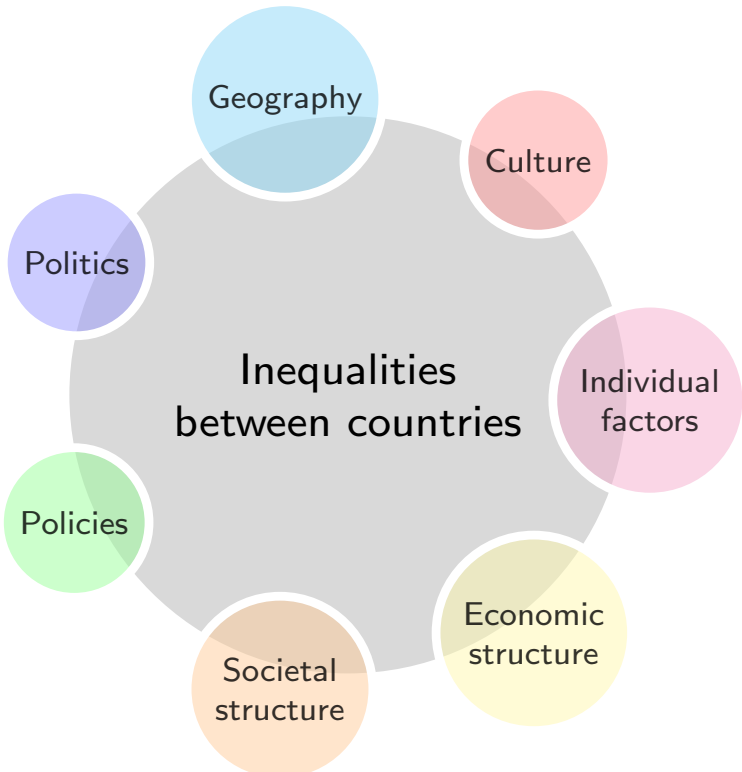
"The **systematic** differences in cancer occurence (incidence, mortality and survival) that exist between and within countries."[1]

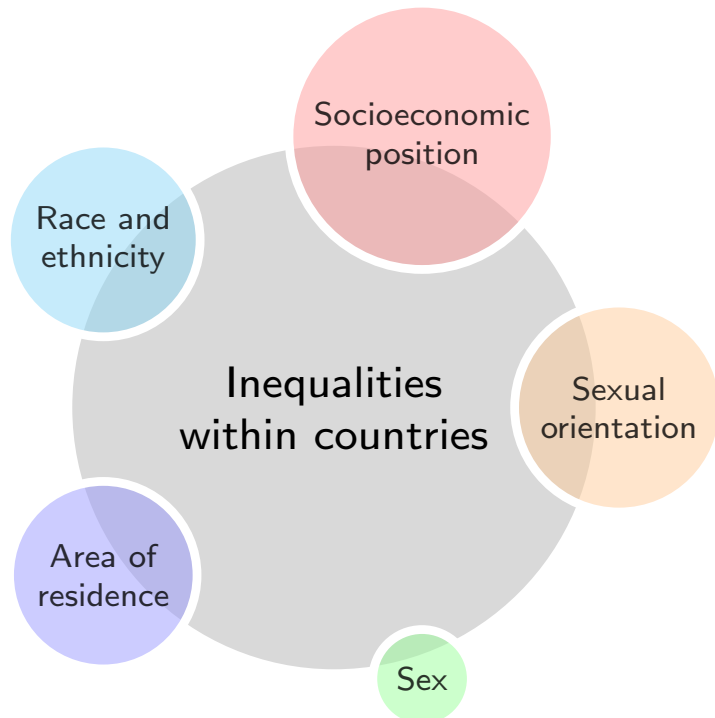
[1] Wild C.P., Weiderpass E., Stewart B.W., editors (2020). *World Cancer Report: Cancer Research for Cancer Prevention.* Lyon, France: International Agency for Research on Cancer. Available from: http://publications.iarc.fr/586. Licence: CC BY-NC-ND 3.0 IGO.

The *unequal* cancer burden

Inequalities in cancer

"The **systematic** differences in cancer occurence (incidence, mortality and survival) that exist between and within countries."[1]



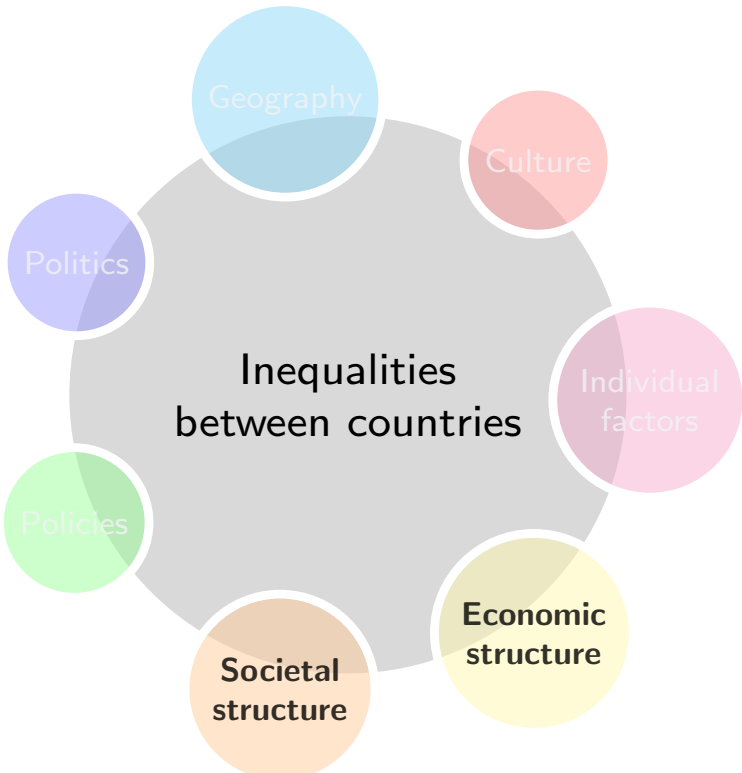


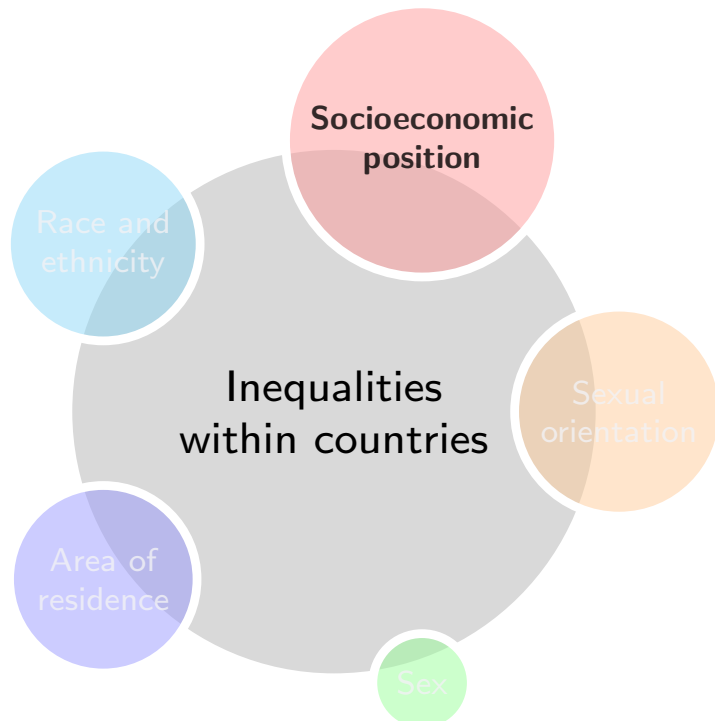
[1] Wild C.P., Weiderpass E., Stewart B.W., editors (2020). *World Cancer Report: Cancer Research for Cancer Prevention.* Lyon, France: International Agency for Research on Cancer. Available from: http://publications.iarc.fr/586. Licence: CC BY-NC-ND 3.0 IGO.

The *unequal* cancer burden

Inequalities in cancer

"The **systematic** differences in cancer occurence (incidence, mortality and survival) that exist between and within countries."[1]

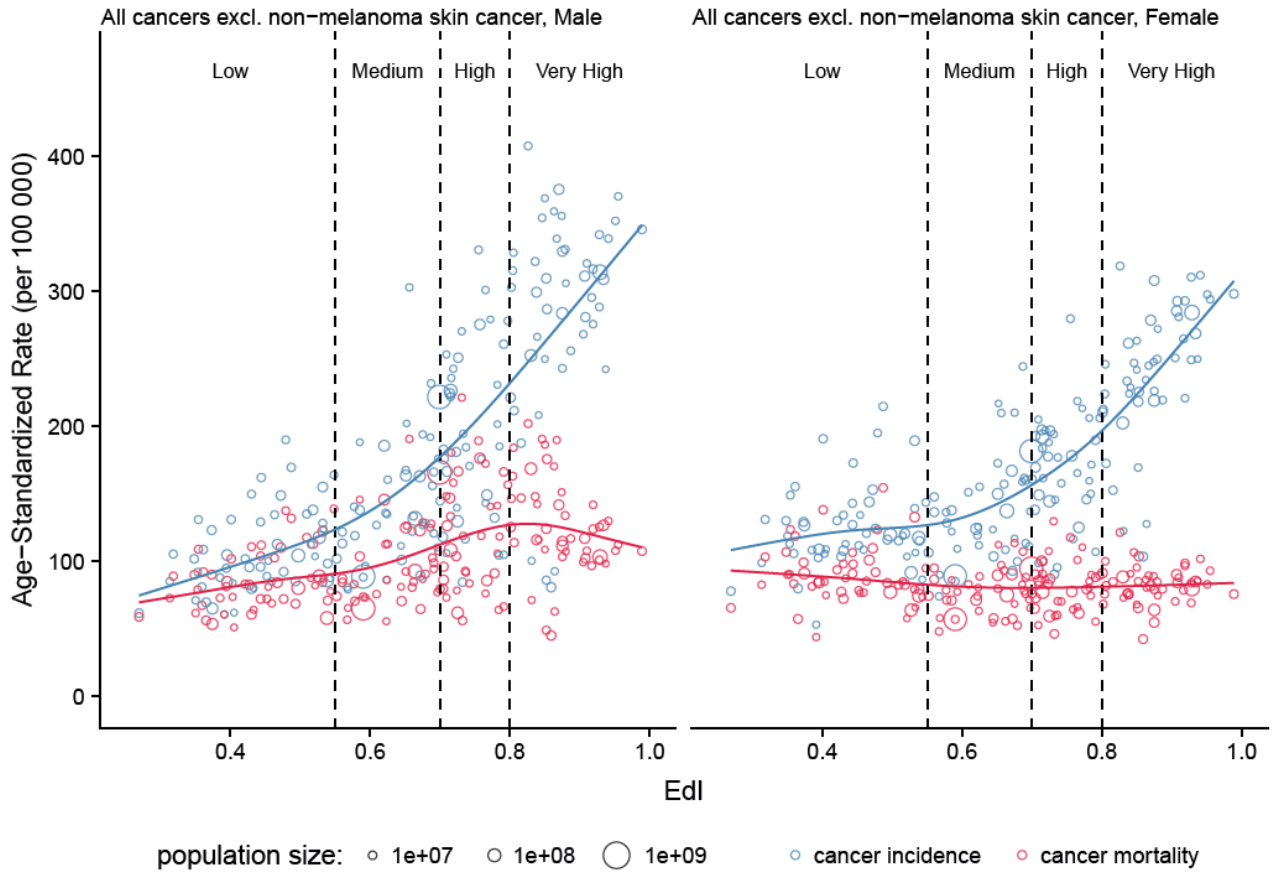




[1] Wild C.P., Weiderpass E., Stewart B.W., editors (2020). *World Cancer Report: Cancer Research for Cancer Prevention.* Lyon, France: International Agency for Research on Cancer. Available from: http://publications.iarc.fr/586. Licence: CC BY-NC-ND 3.0 IGO.

Evidence of cancer inequalities between countries

Evidence of cancer inequalities between countries

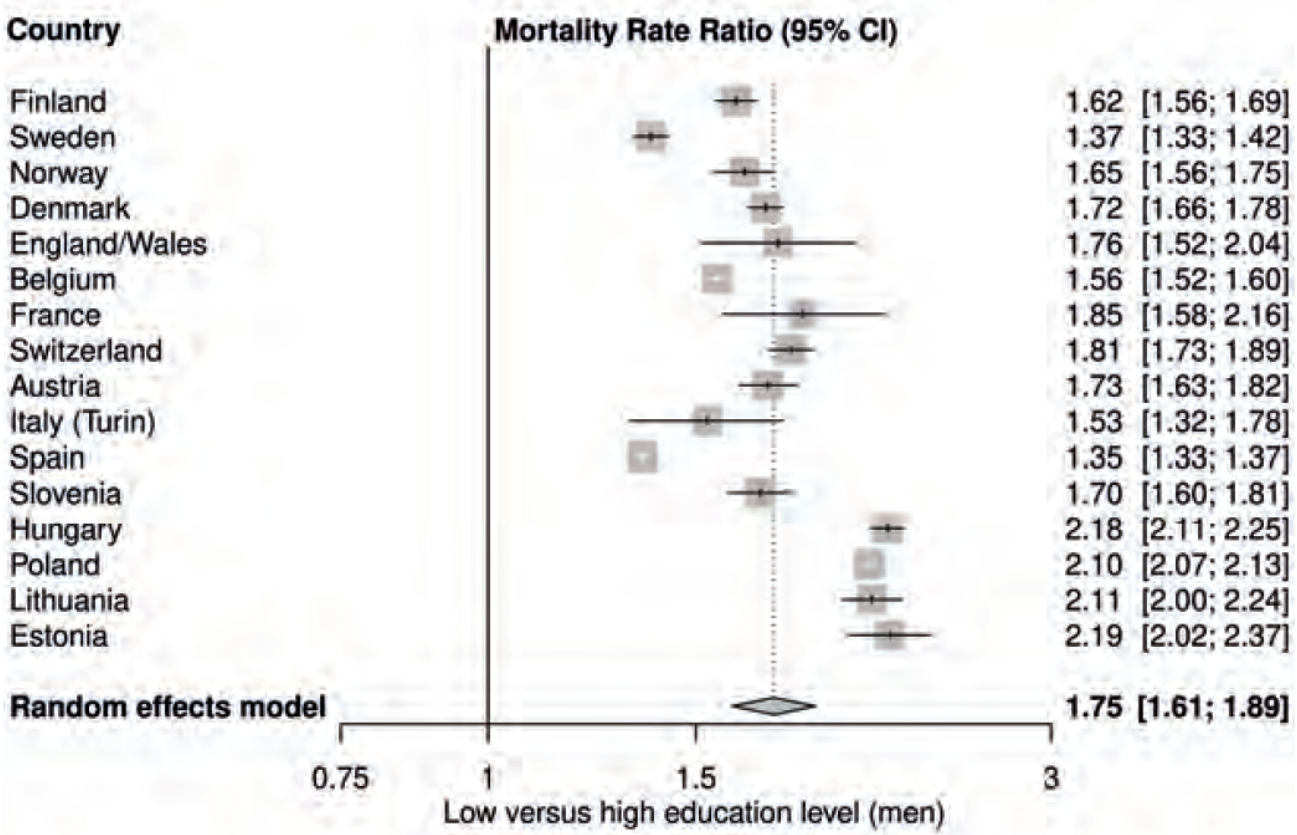


Age-standardize (world population) incidence and mortality rates of all cancer types, by average of socioeconomic development in 2012.

Figure: Wild C.P., Weiderpass E., Stewart B.W., editors (2020). *World Cancer Report: Cancer Research for Cancer Prevention.* Lyon, France: International Agency for Research on Cancer.

Evidence of cancer inequalities within countries

Evidence of cancer inequalities within countries

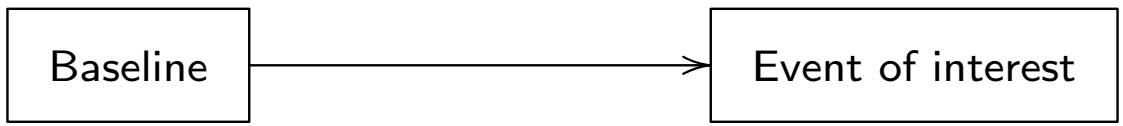


Rate ratios and the corresponding 95% CI of mortality from all cancer combined for men with a low versus high education level

---

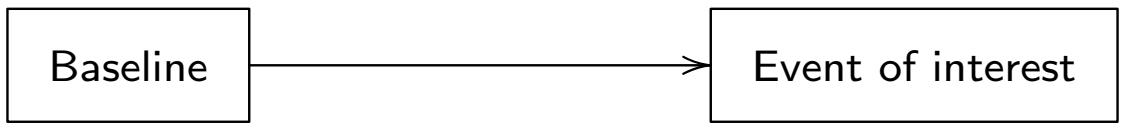
Figure: Wild C.P., Weiderpass E., Stewart B.W., editors (2020). *World Cancer Report: Cancer Research for Cancer Prevention.* Lyon, France: International Agency for Research on Cancer.

Survival analysis



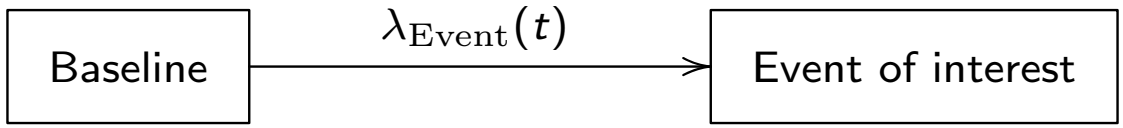
- $T =$ "Time of the event"

Survival analysis



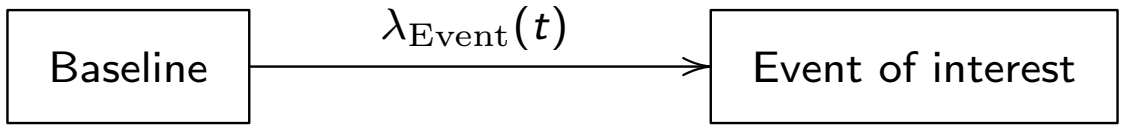
- $T =$ "Time of the event"
- $S(t) = \mathbb{P}(T > t)$ Survival function

Survival analysis



- $T =$ "Time of the event"
- $S(t) = \mathbb{P}(T > t)$ Survival function
- $\lambda(t) = \lim_{dt \to 0} \frac{1}{dt} \mathbb{P}(t \leq T < t + dt \mid T > t)$ Hazard function

Survival analysis

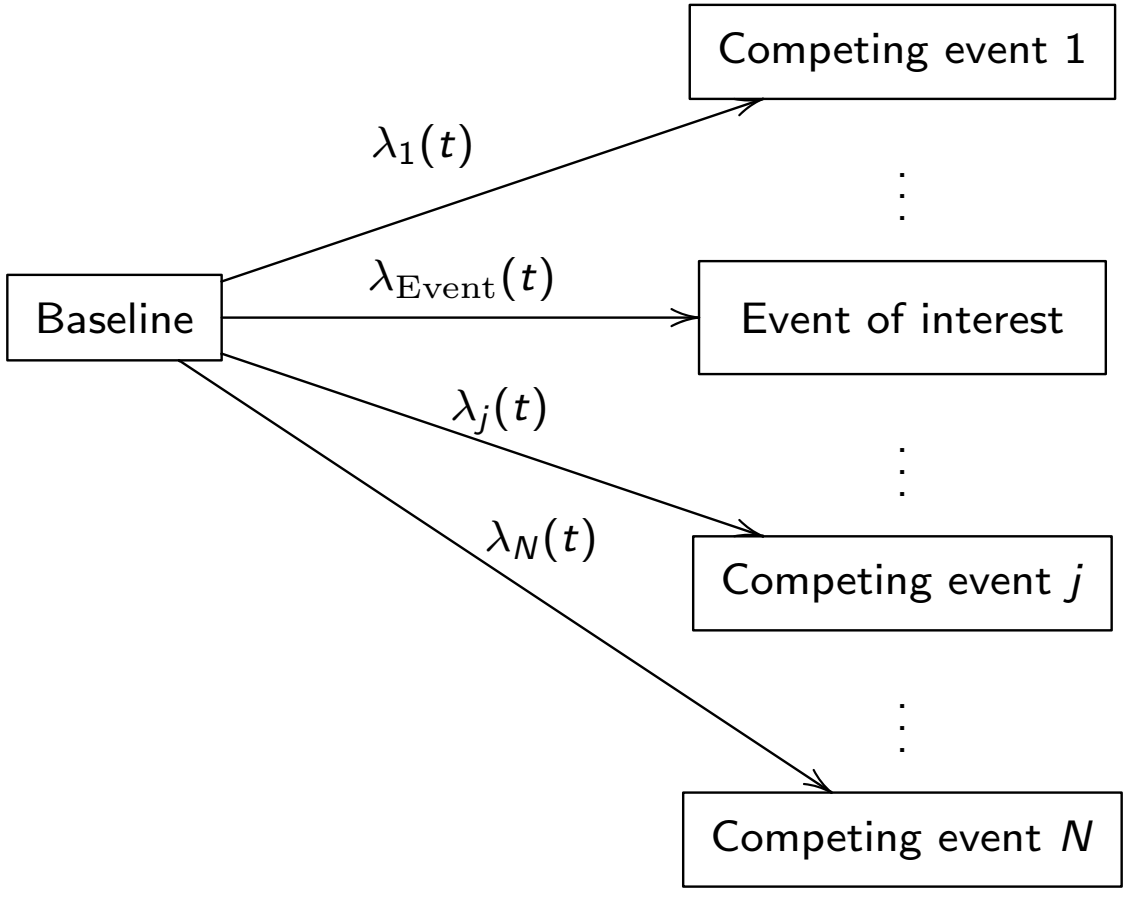


- $T =$ "Time of the event"
- $S(t) = \mathbb{P}(T > t)$ Survival function
- $\lambda(t) = \lim_{dt \to 0} \frac{1}{dt} \mathbb{P}(t \leq T < t + dt \mid T > t)$ Hazard function

Main objective of survival analysis: estimation of $S(t)$ or $\lambda(t)$.

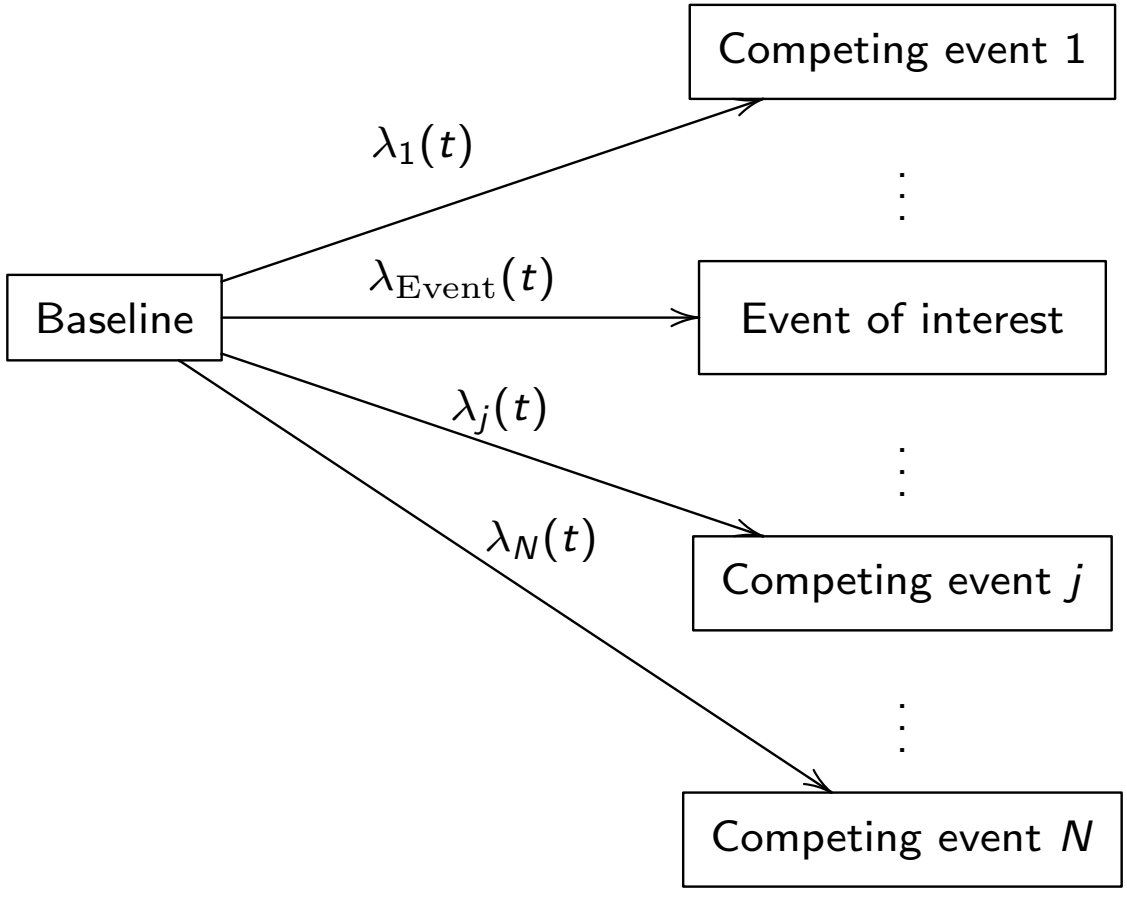
Survival analysis

Life is not easy



Survival analysis

Life is not easy



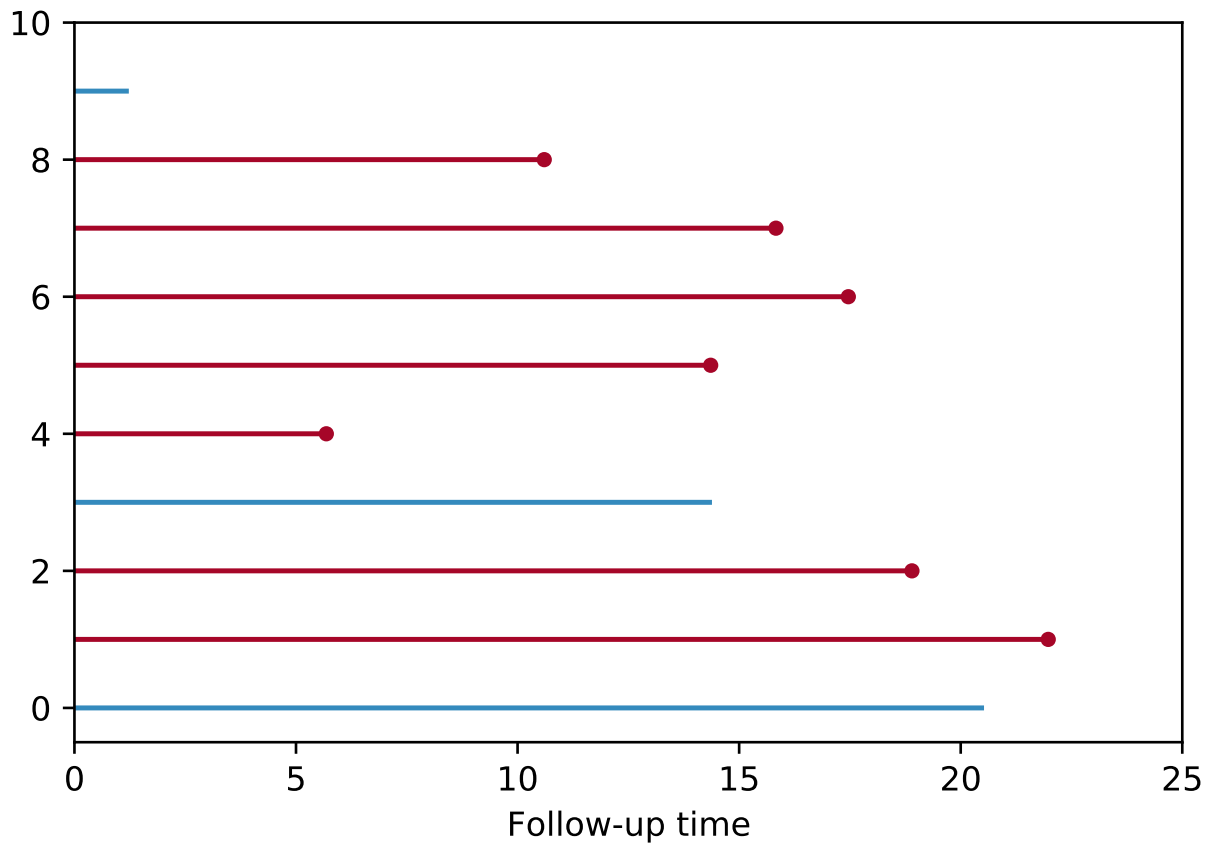
Individuals transitioning to a competing event are informative while they are at risk!

Survival analysis

Censoring

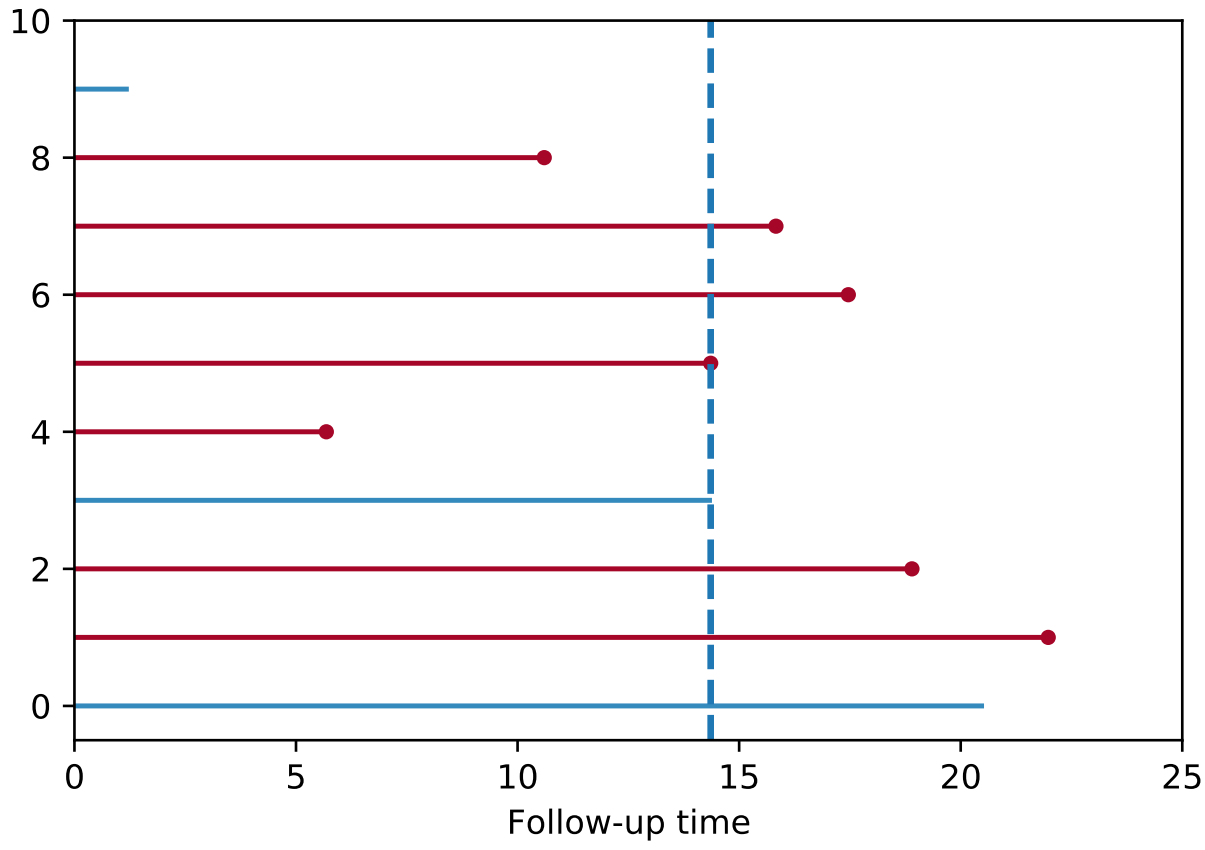
Survival analysis

Censoring



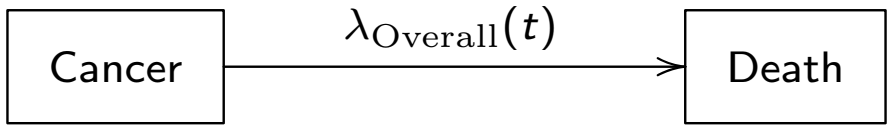
Survival analysis

Censoring

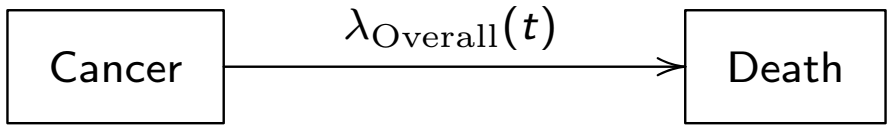


Censored individuals contribute to hazard estimation when they are still at risk

The observed cancer survival

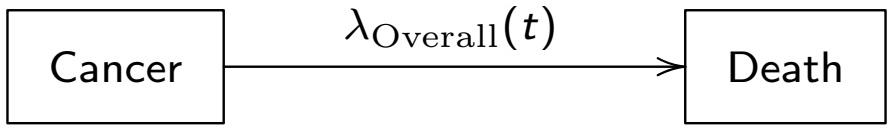


The observed cancer survival



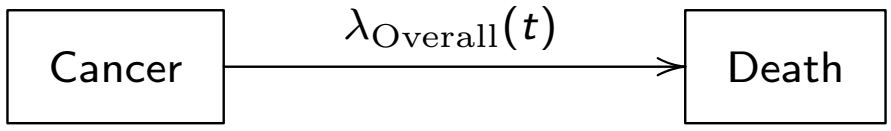
- Not all cancer patients die because of their cancer

The observed cancer survival



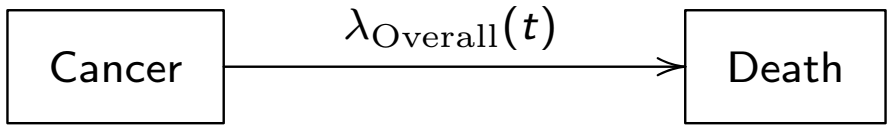
- Not all cancer patients die because of their cancer
- Overall survival is not only determined by cancer

The observed cancer survival



- Not all cancer patients die because of their cancer
- Overall survival is not only determined by cancer
- Genetic, demographic and lifestyle factors depend on general population mortality

The observed cancer survival

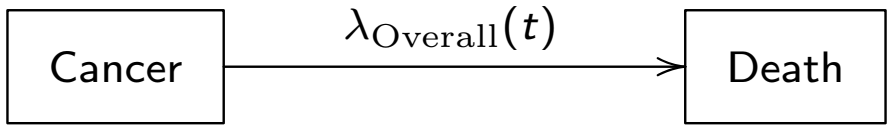


- Not all cancer patients die because of their cancer
- Overall survival is not only determined by cancer
- Genetic, demographic and lifestyle factors depend on general population mortality

When we want to

- Test the efficiency of health-care systems

The observed cancer survival

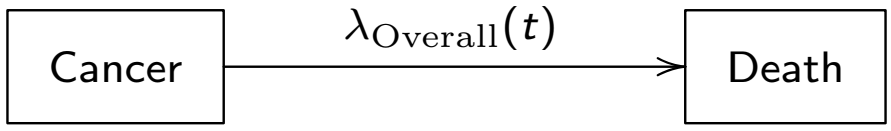


- Not all cancer patients die because of their cancer
- Overall survival is not only determined by cancer
- Genetic, demographic and lifestyle factors depend on general population mortality

When we want to

- Test the efficiency of health-care systems
- Compare cancer survival among different countries or periods

The observed cancer survival

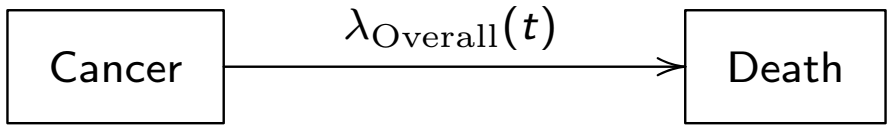


- Not all cancer patients die because of their cancer
- Overall survival is not only determined by cancer
- Genetic, demographic and lifestyle factors depend on general population mortality

When we want to

- Test the efficiency of health-care systems
- Compare cancer survival among different countries or periods
- Measure the cancer burden among different populations

The observed cancer survival



- Not all cancer patients die because of their cancer
- Overall survival is not only determined by cancer
- Genetic, demographic and lifestyle factors depend on general population mortality

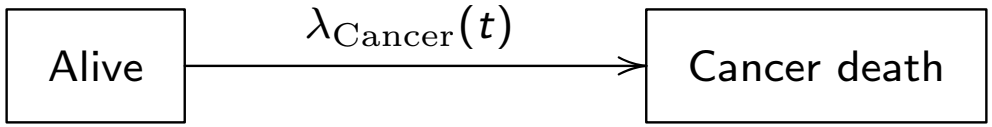
When we want to

- Test the efficiency of health-care systems
- Compare cancer survival among different countries or periods
- Measure the cancer burden among different populations

we need a mortality indicator which is independent of the general population mortality

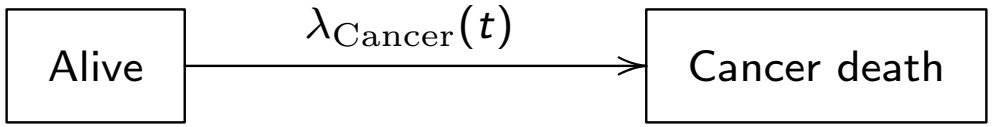
The concept of net survival

Net survival is the survival that would be observed in a **hypothetical world** where cancer would be the only cause of death



The concept of net survival

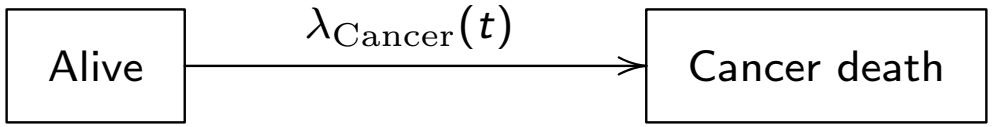
Net survival is the survival that would be observed in a **hypothetical world** where cancer would be the only cause of death



- Clinically denied of sense

The concept of net survival

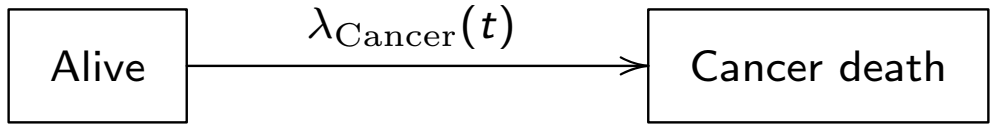
Net survival is the survival that would be observed in a **hypothetical world** where cancer would be the only cause of death



- Clinically denied of sense
- The only indicator allowing to measure the effect of cancer independently of the other causes of death

The concept of net survival

Net survival is the survival that would be observed in a **hypothetical world** where cancer would be the only cause of death



- Clinically denied of sense
- The only indicator allowing to measure the effect of cancer independently of the other causes of death

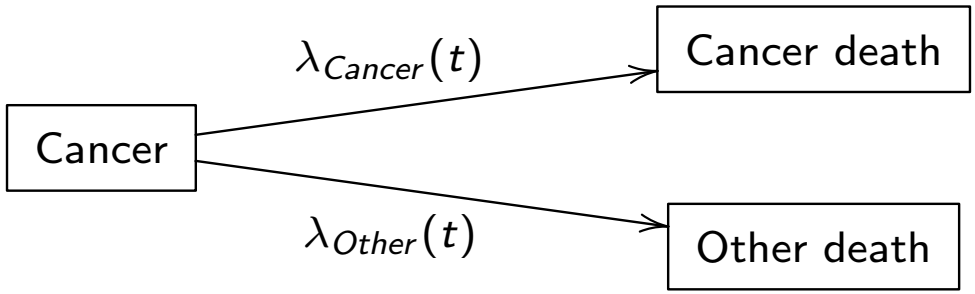
How to estimate survival in a **hypothetical** world?

Cause-specific survival

If causes of death are available:

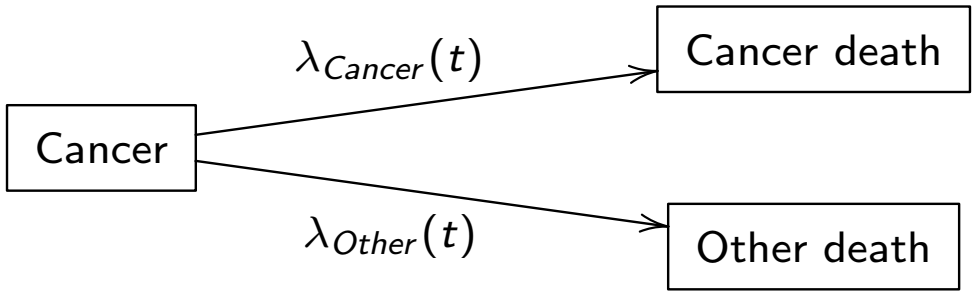
Cause-specific survival

If causes of death are available:



Cause-specific survival

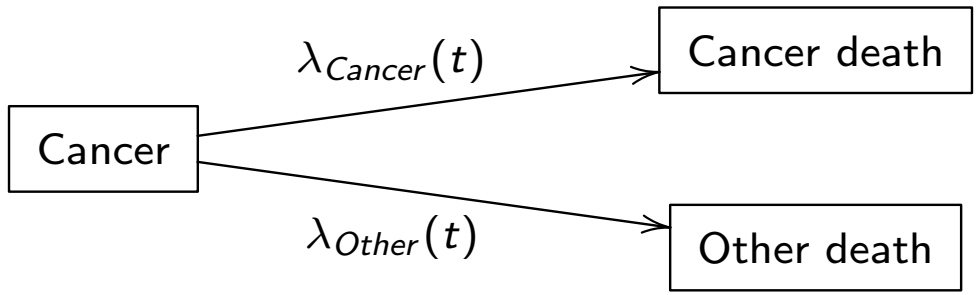
If causes of death are available:



Main sources of bias

Cause-specific survival

If causes of death are available:

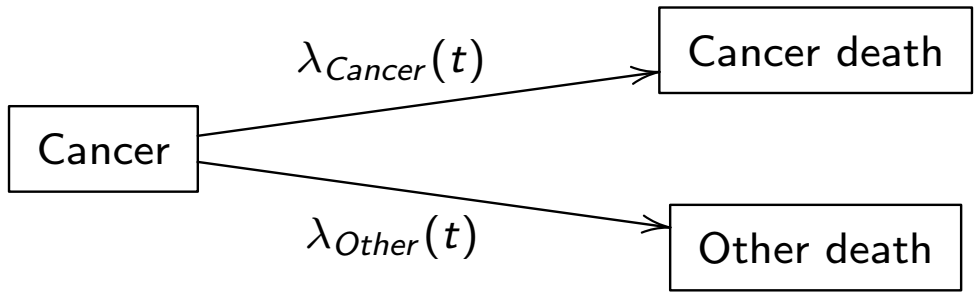


Main sources of bias

- Causes of death **misclassification**

Cause-specific survival

If causes of death are available:

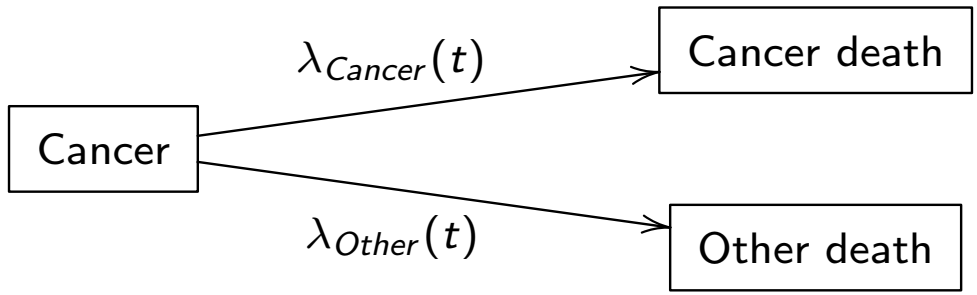


Main sources of bias

- Causes of death **misclassification**
  - Lack of sensitivity/specificity

Cause-specific survival

If causes of death are available:

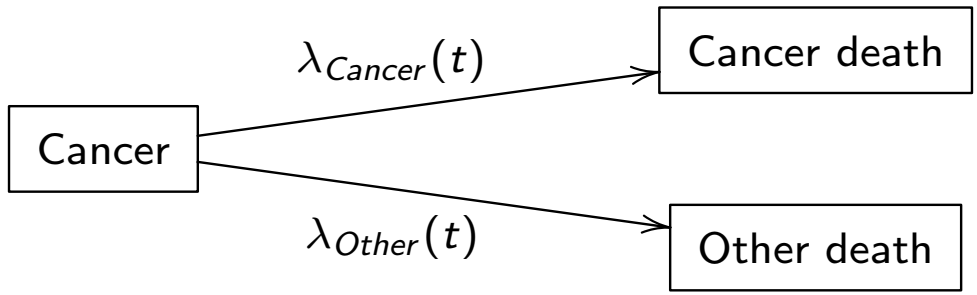


Main sources of bias

- Causes of death **misclassification**
  - Lack of sensitivity/specificity
  - Identification of the **underlying cause of death**

Cause-specific survival

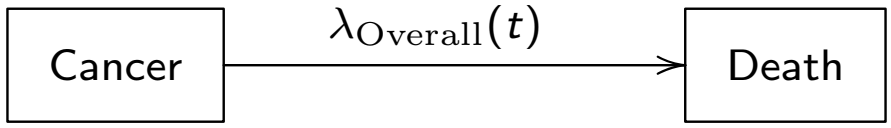
If causes of death are available:



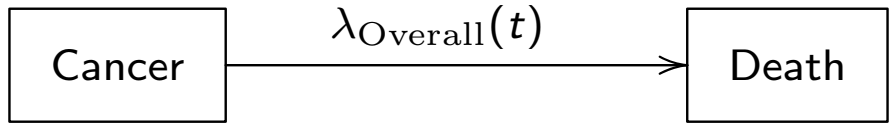
Main sources of bias

- Causes of death **misclassification**
  - Lack of sensitivity/specificity
  - Identification of the **underlying cause of death**
- Causes of death are not always available and/or reliable

Relative survival



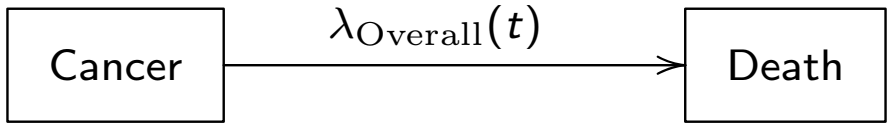
Relative survival



$$\lambda_{\mathrm{Overall}}(t; x, z, a) = \lambda_{\mathrm{Cancer}}(t; x, a) + \lambda_{\mathrm{Other}}(a + t; z)$$

a: age at diagnosis, x, z: set of covariates

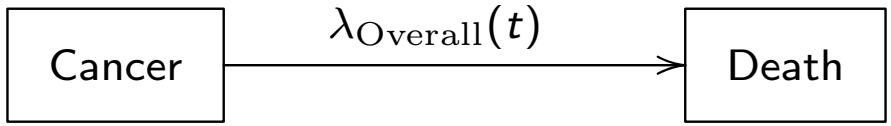
Relative survival



$$\underbrace{\boxed{\lambda_{\text{Overall}}(t; x, z, a)}}_{\text{Cancer cohort}} = \lambda_{\text{Cancer}}(t; x, a) + \lambda_{\text{Other}}(a + t; z)$$

a: age at diagnosis, x, z: set of covariates

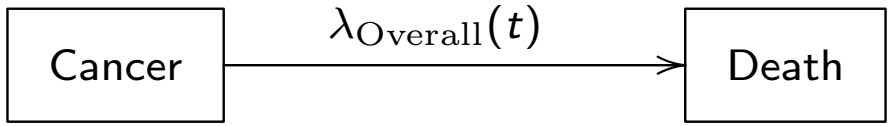
Relative survival



$$\underbrace{\boxed{\lambda_{\text{Overall}}(t; x, z, a)}}_{\text{Cancer cohort}} = \lambda_{\text{Cancer}}(t; x, a) + \underbrace{\boxed{\lambda_{\text{Other}}(a + t; z)}}_{\text{General population}}$$

a: age at diagnosis, x, z: set of covariates

Relative survival



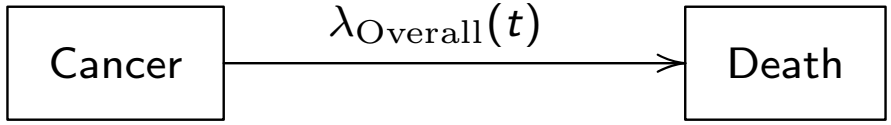
$$\underbrace{\boxed{\lambda_{\text{Overall}}(t; x, z, a)}}_{\text{Cancer cohort}} = \lambda_{\text{Cancer}}(t; x, a) + \underbrace{\boxed{\lambda_{\text{Other}}(a + t; z)}}_{\text{General population}}$$

a: age at diagnosis, x, z: set of covariates

Main source of bias

- **Lack of comparability** between cancer cohort and general population

Relative survival



$$\underbrace{\boxed{\lambda_{\text{Overall}}(t; x, z, a)}}_{\text{Cancer cohort}} = \lambda_{\text{Cancer}}(t; x, a) + \underbrace{\boxed{\lambda_{\text{Other}}(a + t; z)}}_{\text{General population}}$$

a: age at diagnosis, x, z: set of covariates

Main source of bias

- **Lack of comparability** between cancer cohort and general population

False in general!

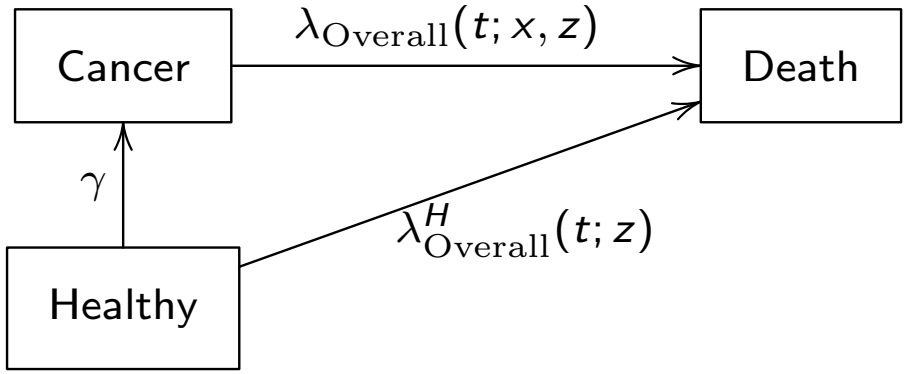
$$\lambda_{\text{Other}}^{\text{Cancer}} \neq \lambda_{\text{Other}}^{\text{General population}}$$

Relative survival

When it is possible to estimate λ_{Other}

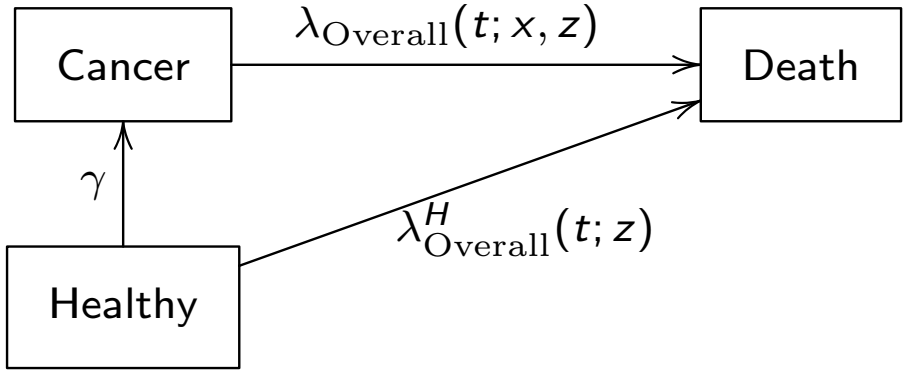
Relative survival

When it is possible to estimate λ_{Other}



Relative survival

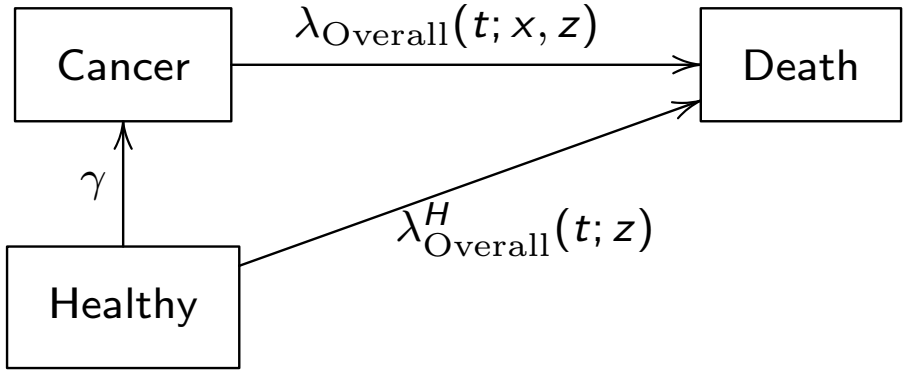
When it is possible to estimate λ_{Other}



Comparability hypothesis: $\lambda^{H}_{\mathrm{Overall}}(t; z) \approx \lambda_{\mathrm{Other}}(t; z)$ for all $t \geq 0$.

Relative survival

When it is possible to estimate λ_{Other}

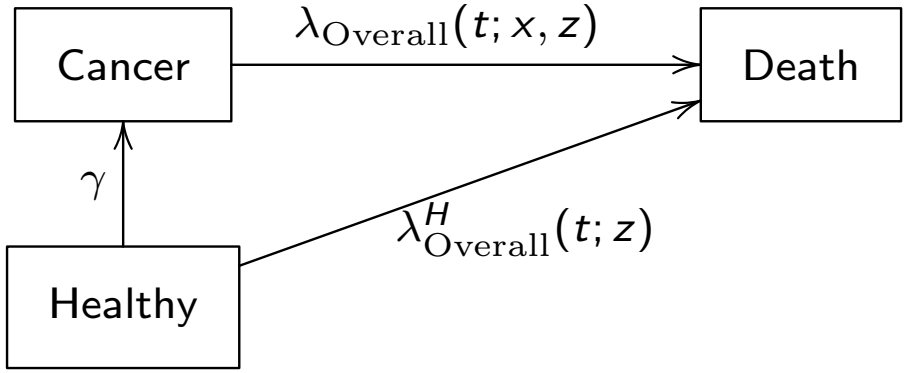


Comparability hypothesis: $\lambda^{H}_{\mathrm{Overall}}(t; z) \approx \lambda_{\mathrm{Other}}(t; z)$ for all $t \geq 0$.

Relative survival estimation: $\lambda_{\mathrm{Overall}}(t; x, z) = \lambda_{\mathrm{Cancer}}(t; x) + \lambda^{H}_{\mathrm{Overall}}(t; z)$

Relative survival

When it is possible to estimate λ_{Other}



Comparability hypothesis: $\lambda^{H}_{\mathrm{Overall}}(t; z) \approx \lambda_{\mathrm{Other}}(t; z)$ for all $t \geq 0$.

Relative survival estimation: $\lambda_{\mathrm{Overall}}(t; x, z) = \lambda_{\mathrm{Cancer}}(t; x) + \lambda^{H}_{\mathrm{Overall}}(t; z)$

Acess to two comparable populations is unsual/difficult

Net survival estimation on EPIC cohort

Javier González-Delgado, Vivian Viallon[3], Grégoire Rey[4] and Hadrien Charvat[5]

3. Nutritional Methodology and Biostatistics Branch, International Agency for Research on Cancer, Lyon, France.
4 Inserm, Centre d'épidémiologie sur les causes médicales de décès (CépiDc), Le Kremlin-Bicêtre.
5. Cancer Surveillance Branch, International Agency for Research on Cancer, Lyon, France.

Net survival estimation on EPIC cohort

Javier González-Delgado, Vivian Viallon[3], Grégoire Rey[4] and Hadrien Charvat[5]

3. Nutritional Methodology and Biostatistics Branch, International Agency for Research on Cancer, Lyon, France.
4 Inserm, Centre d'épidémiologie sur les causes médicales de décès (CépiDc), Le Kremlin-Bicêtre.
5. Cancer Surveillance Branch, International Agency for Research on Cancer, Lyon, France.

European Prospective Investigation into Cancer and Nutrition (EPIC)[2]

- ∼521.000 participants recruited across 23 European centers
- Followed for almost 15 years
- More than 58.000 reported deaths
- 75 covariates describing dietary exposure, lifestyle factors, anthropometry and biological parameters
- Acess to high quality cause of death information
- 67.000 participants diagnosed with cancer

[2]International Agency for Research on Cancer. EPIC study. 2020. Retrieved from https://epic.iarc.fr

Net survival estimation on EPIC cohort

General population and EPIC cohort are not comparable

Net survival estimation on EPIC cohort

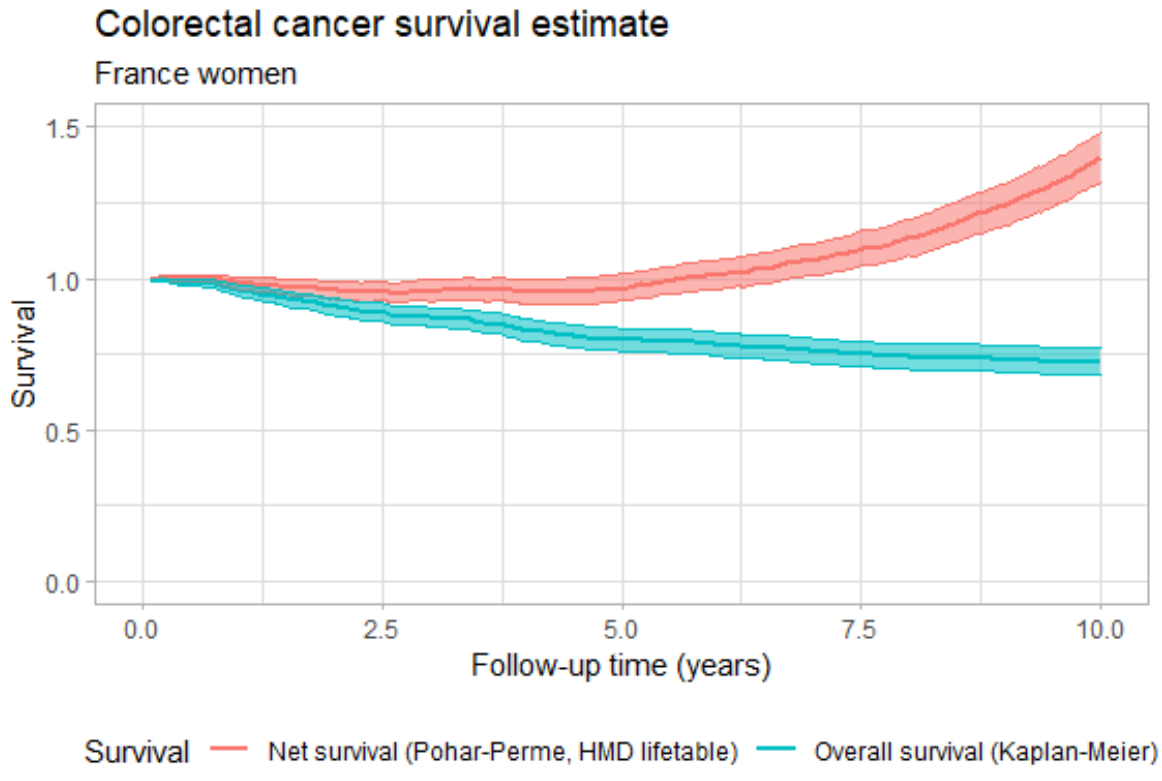
General population and EPIC cohort are not comparable

EPIC relative survival estimate using general population mortality to estimate λ_{Other}.

Net survival estimation on EPIC cohort

General population and EPIC cohort are not comparable

EPIC relative survival estimate using general population mortality to estimate λ_{Other}.



Net survival estimation on EPIC cohort

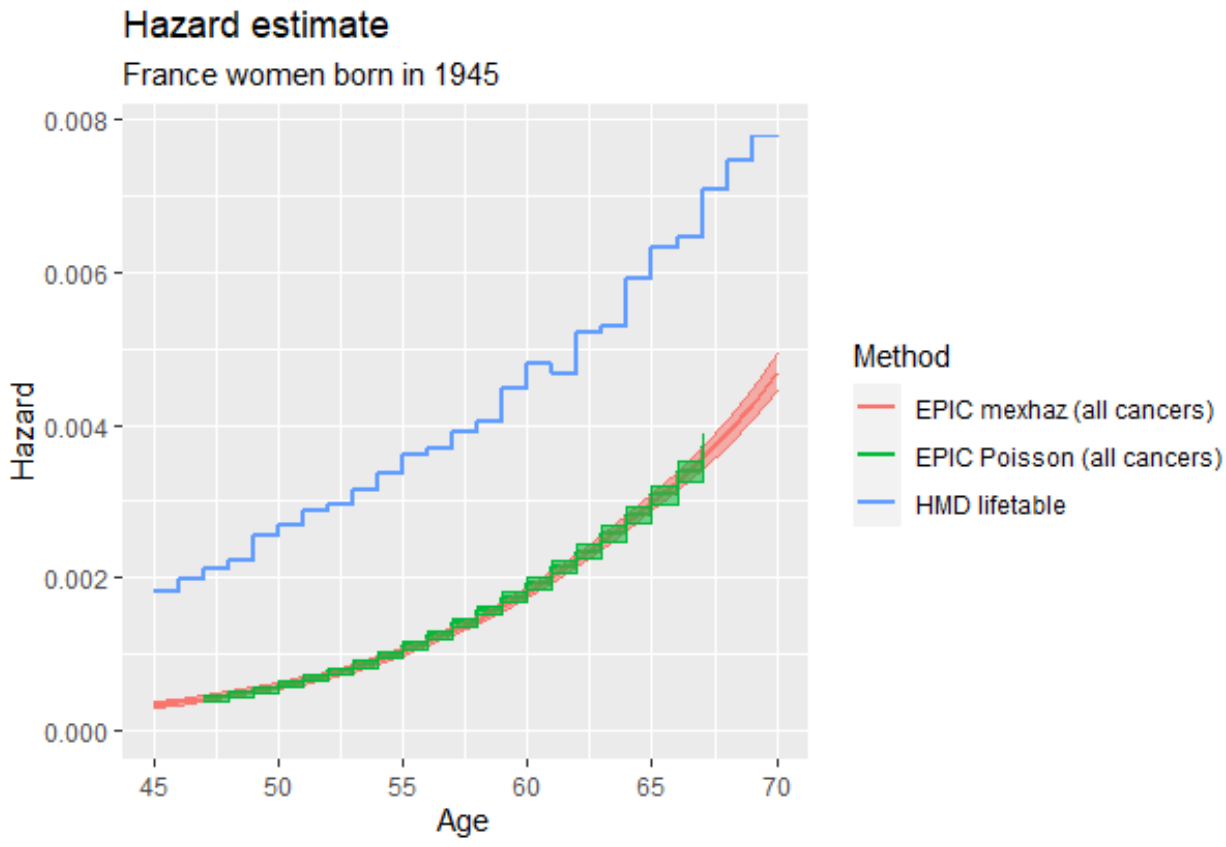
General population and EPIC cohort are not comparable

EPIC λ_{Other} and general population hazard estimates

Net survival estimation on EPIC cohort

General population and EPIC cohort are not comparable

EPIC λ_{Other} and general population hazard estimates

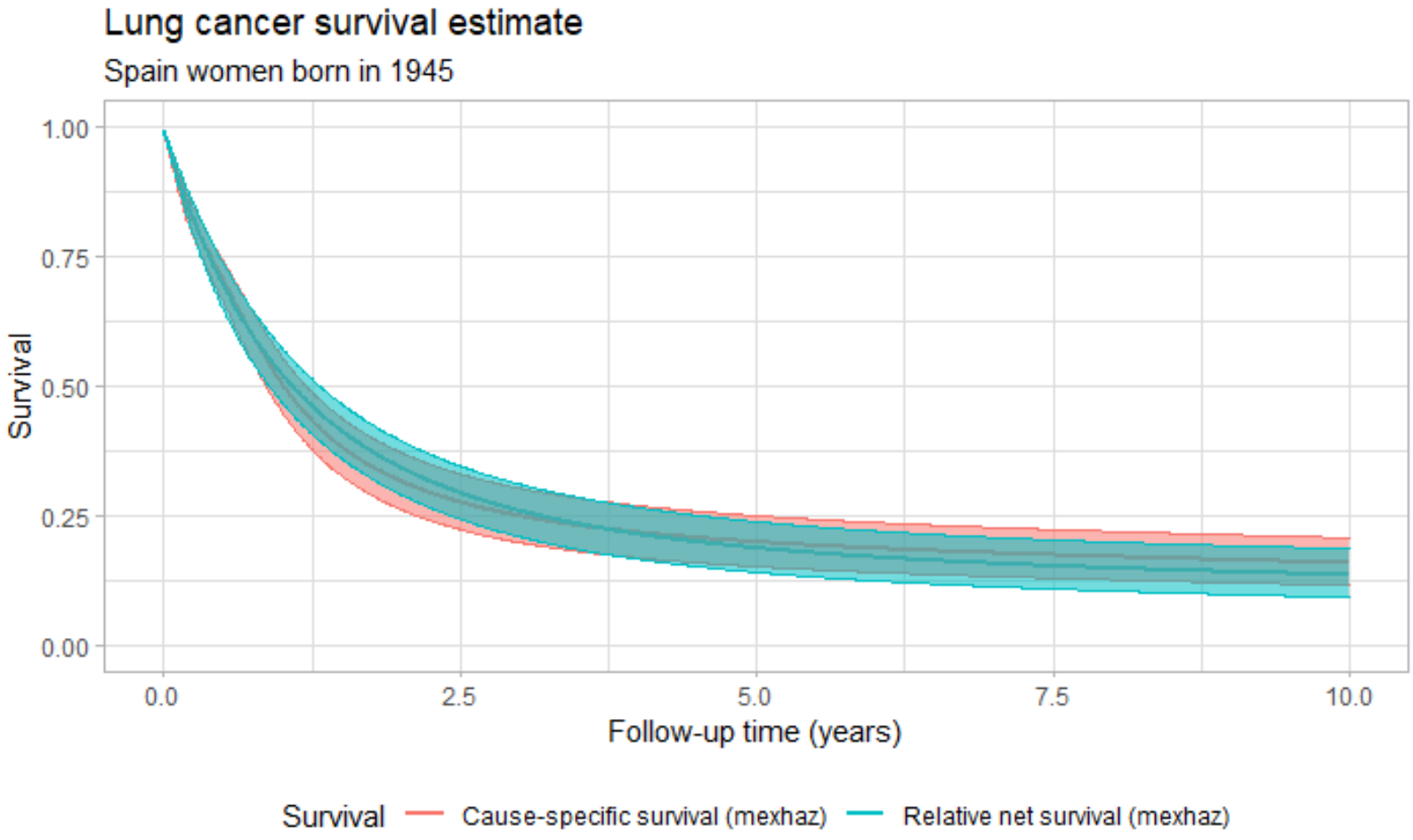


Net survival estimation on EPIC cohort

Relative and cause-specific survival

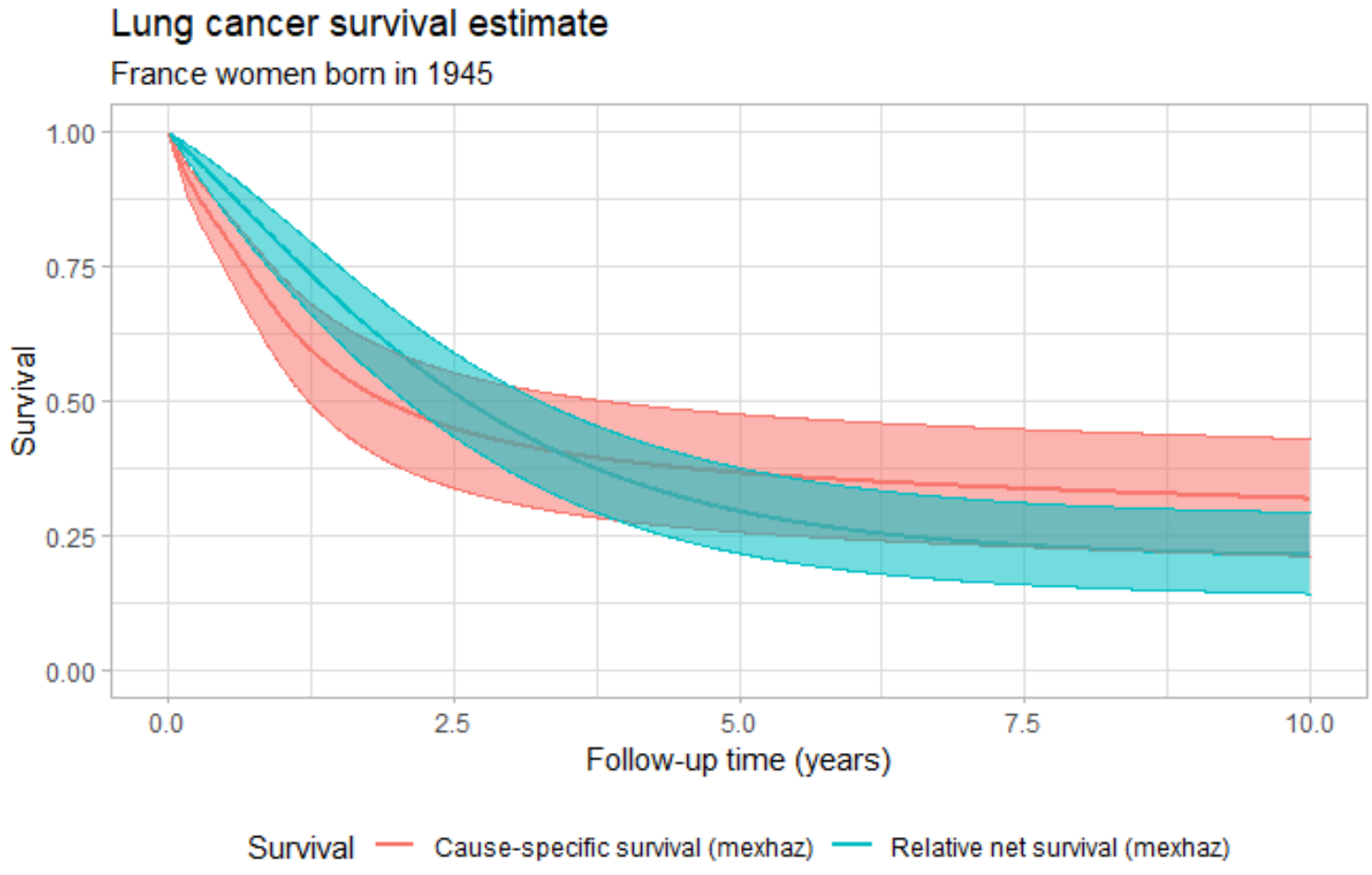
Net survival estimation on EPIC cohort

Relative and cause-specific survival



Net survival estimation on EPIC cohort

Relative and cause-specific survival

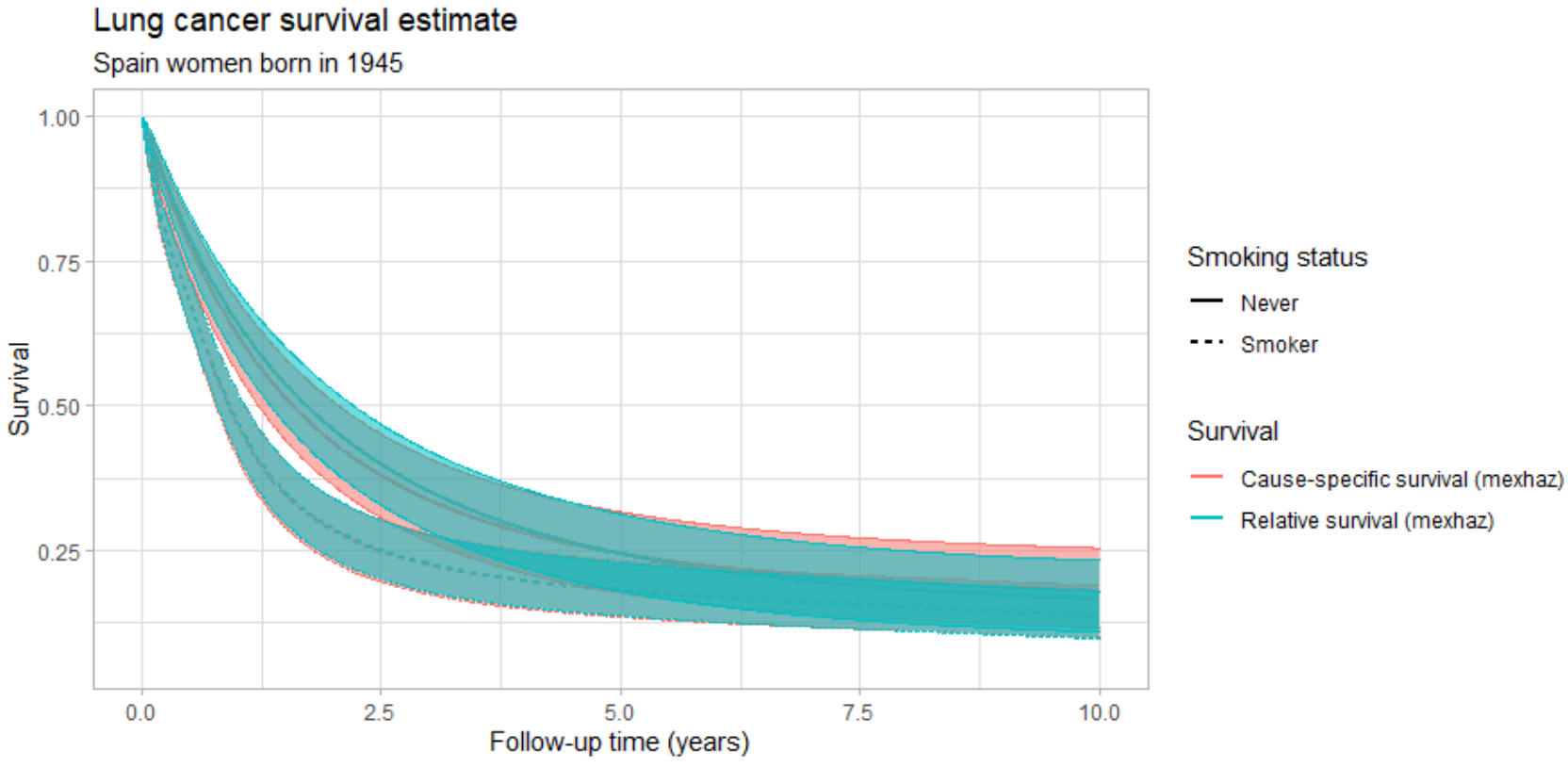


Net survival estimation on EPIC cohort

Relative and cause-specific survival: taking covariates into account

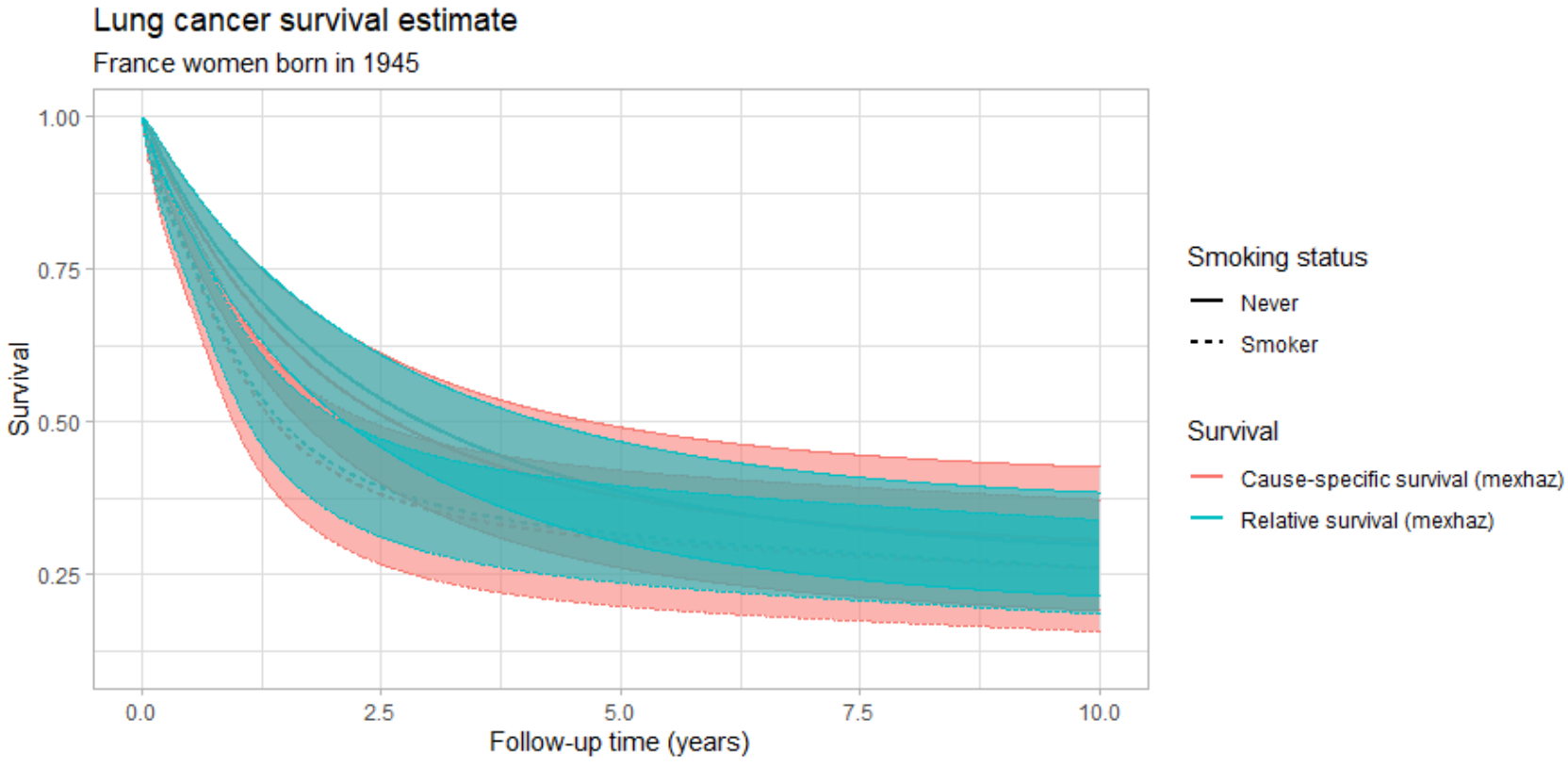
Net survival estimation on EPIC cohort

Relative and cause-specific survival: taking covariates into account



Net survival estimation on EPIC cohort

Relative and cause-specific survival: taking covariates into account



Some conclusions

- Relative and cause-specific survival curves are **superimposed or coincident within the CI**.

Some conclusions

- Relative and cause-specific survival curves are **superimposed or coincident within the CI**.
- **Adjusting by smoking status** may correct differences for **smoker** patients.

Some conclusions

- Relative and cause-specific survival curves are **superimposed or coincident within the CI**.
- **Adjusting by smoking status** may correct differences for **smoker** patients.
- **Cause-specific** method **is recommended** due to its simplicity.

Some conclusions

- Relative and cause-specific survival curves are **superimposed or coincident within the CI**.
- **Adjusting by smoking status** may correct differences for **smoker** patients.
- **Cause-specific** method **is recommended** due to its simplicity.
- But... results can **not** be **generalized** to general population.

Some conclusions

- Relative and cause-specific survival curves are **superimposed or coincident within the CI**.
- **Adjusting by smoking status** may correct differences for **smoker** patients.
- **Cause-specific** method **is recommended** due to its simplicity.
- But... results can **not** be **generalized** to general population.

Thank you for your attention!