When survival bias is not truly a bias, but rather the correct answer to the wrong question

L. Paloma Rojas-Saunero MD, PhD
Postdoctoral scholar
Department of Epidemiology, UCLA

UCLA Fielding School of Public Health

Society for Epidemiologic Research

Competing events

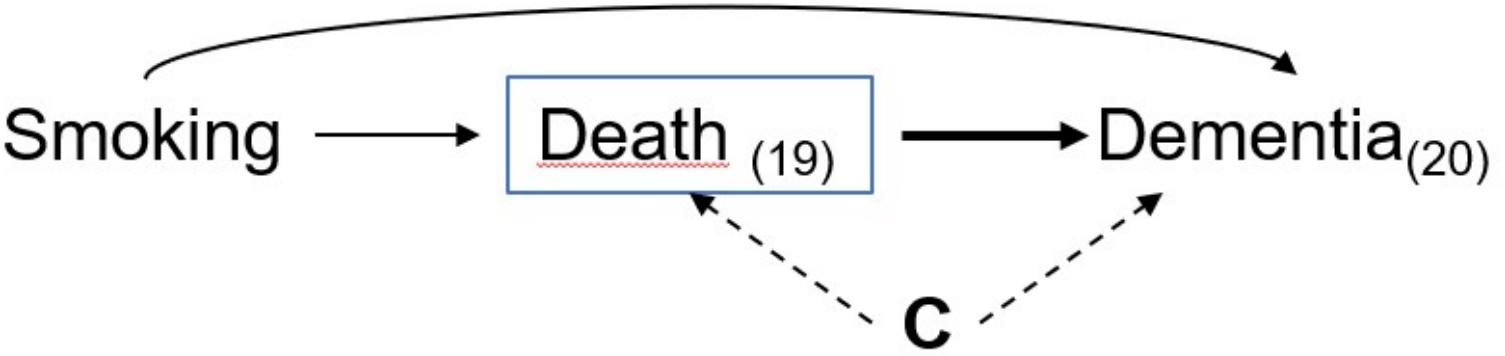
Events that *preclude** the outcome of interest.

Competing events

Events that *preclude** the outcome of interest.

Competing events

Events that *preclude** the outcome of interest.



Causal estimands

- Contrast of (counterfactual) outcome distributions in the **same individuals** but under **different levels of exposure**.
- The only explanation for a difference is the exposure, not comparing different individuals.

Total effect

Total effect

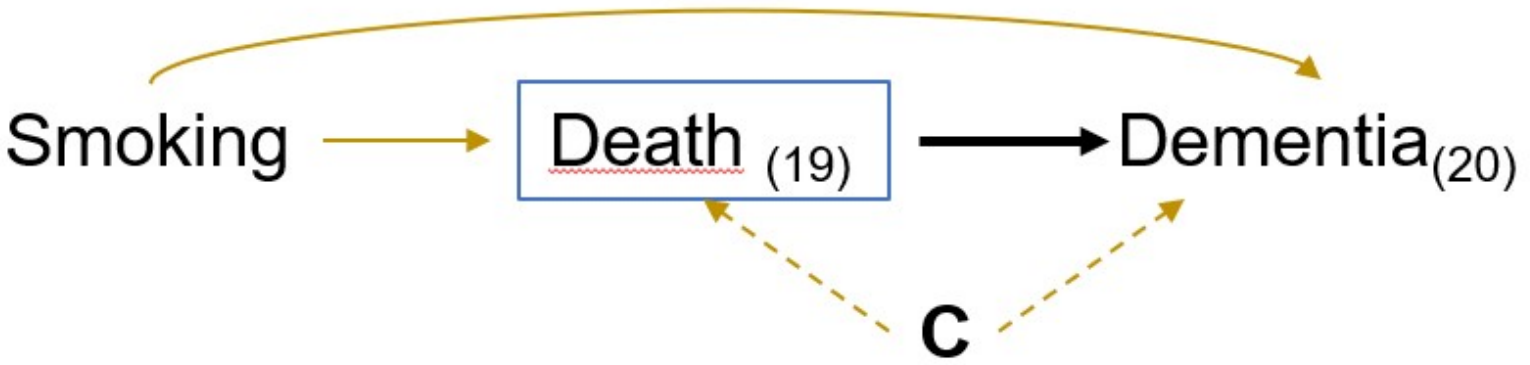
What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking, compared to had all individuals continued smoking?

$$Pr[Y_{20}^{a=1} = 1] - Pr[Y_{20}^{a=0} = 1]$$

Total effect

What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking, compared to had all individuals continued smoking?

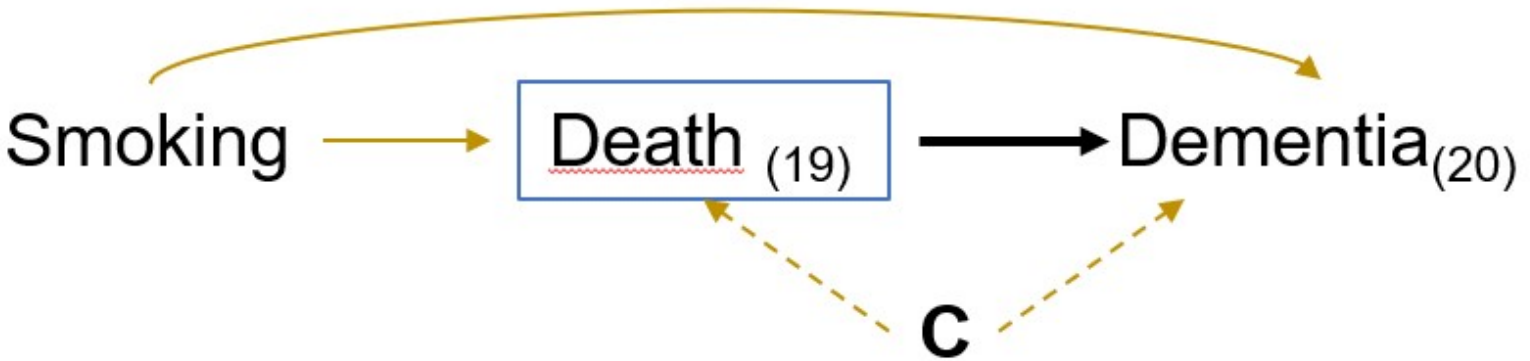
$$Pr[Y_{20}^{a=1} = 1] - Pr[Y_{20}^{a=0} = 1]$$



Total effect

What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking, compared to had all individuals continued smoking?

$$Pr[Y_{20}^{a=1} = 1] - Pr[Y_{20}^{a=0} = 1]$$



** *Cause-specific cumulative incidence* or *crude* risk

Controlled direct effect

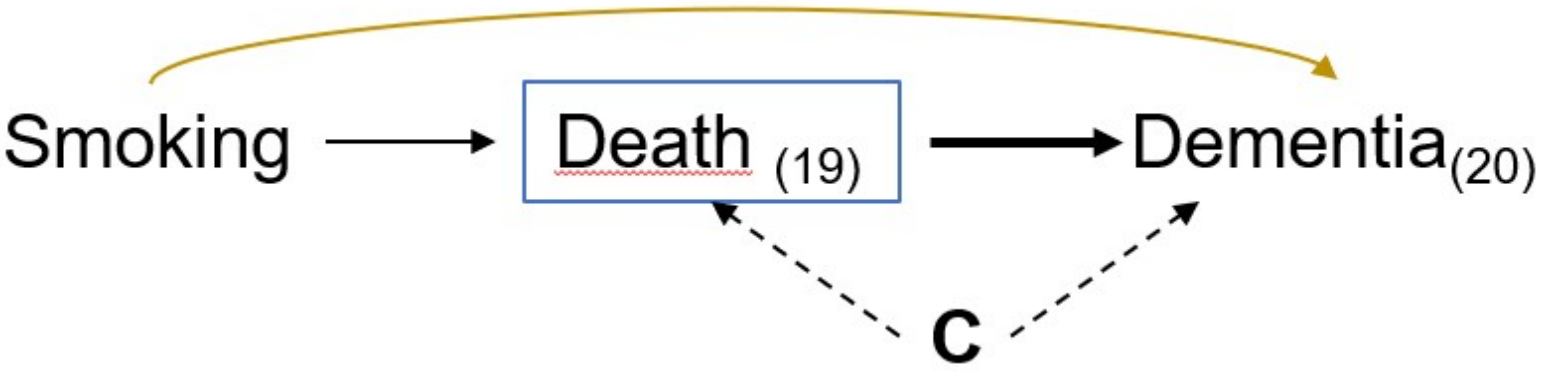
What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking *and not died* during the study period, compared to had all individuals continued smoking *and not died* ?

$$Pr[Y_{20}^{a=1,d_{19}=0} = 1] - Pr[Y_{20}^{a=0,d_{19}=0} = 1]$$

Controlled direct effect

What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking *and not died* during the study period, compared to had all individuals continued smoking *and not died* ?

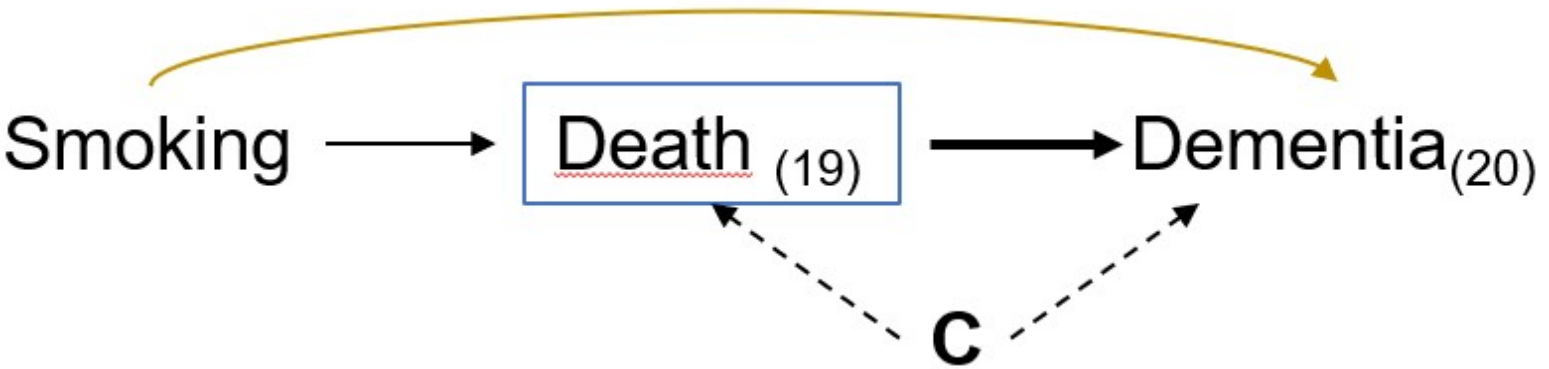
$$Pr[Y_{20}^{a=1,d_{19}=0} = 1] - Pr[Y_{20}^{a=0,d_{19}=0} = 1]$$



Controlled direct effect

What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking *and not died* during the study period, compared to had all individuals continued smoking *and not died* ?

$$Pr[Y_{20}^{a=1,d_{19}=0} = 1] - Pr[Y_{20}^{a=0,d_{19}=0} = 1]$$

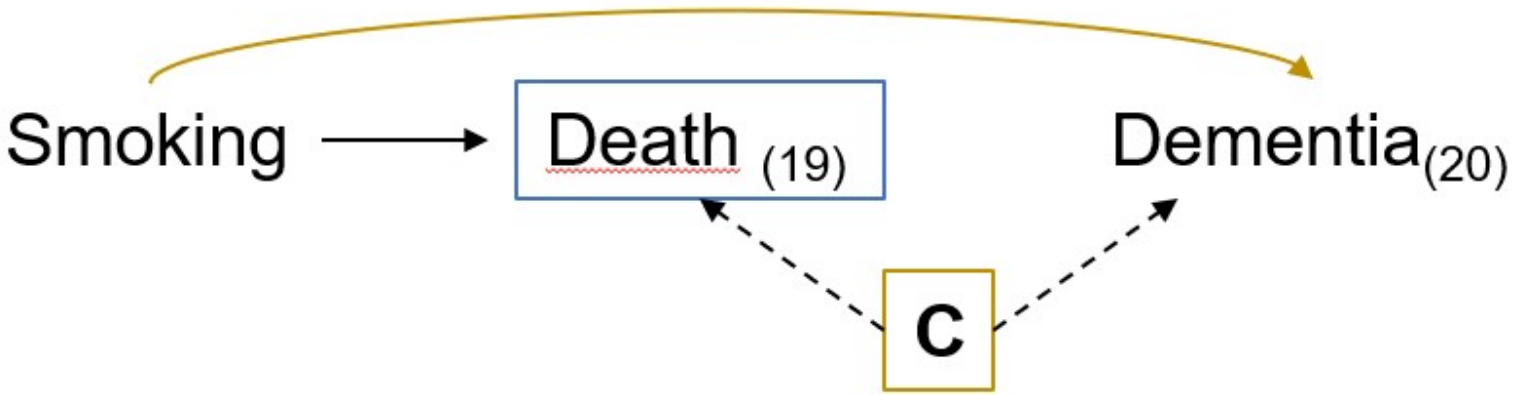


** *Marginal* or *net* risk

Controlled direct effect

What is the risk** of dementia at 20 years of follow-up had all individuals stopped smoking *and not died* during the study period, compared to had all individuals continued smoking *and not died* ?

$$Pr[Y_{20}^{a=1,d_{19}=0} = 1] - Pr[Y_{20}^{a=0,d_{19}=0} = 1]$$



** *Marginal* or *net* risk

Identifiability assumptions for death (competing event)

Assumption	Total Effect	Controlled direct effect
Exchangeability	Not needed	Death is independent of future outcomes had everyone followed A = a and death was eliminated, conditional on covariates

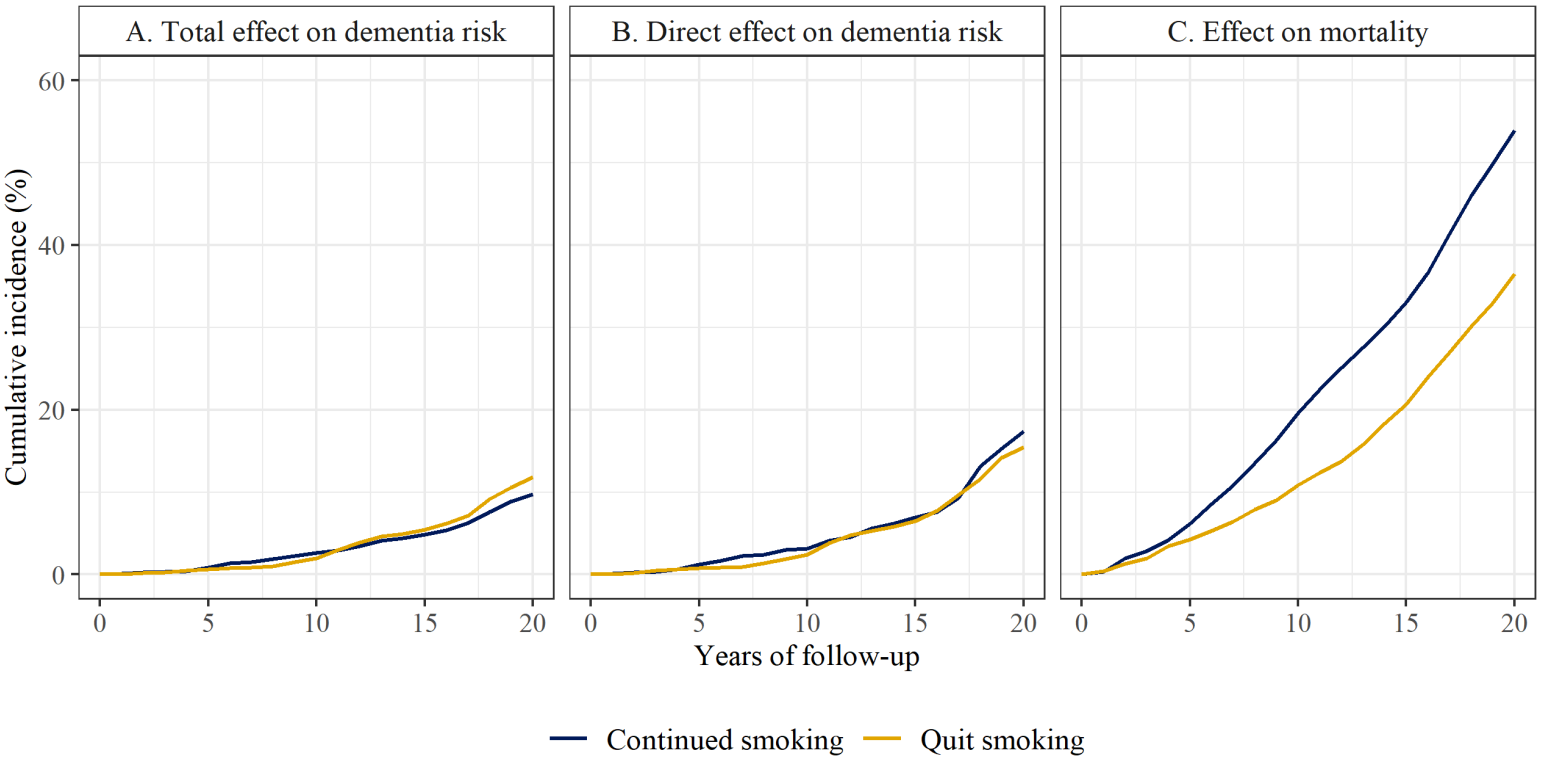
Identifiability assumptions for death (competing event)

Assumption	Total Effect	Controlled direct effect
Exchangeability	Not needed	Death is independent of future outcomes had everyone followed A = a and death was eliminated, conditional on covariates
Positivity	Not needed	At every follow-up time, there are individuals with any possibly observed level A = a and covariate history who remain alive and free of dementia diagnosis.

Identifiability assumptions for death (competing event)

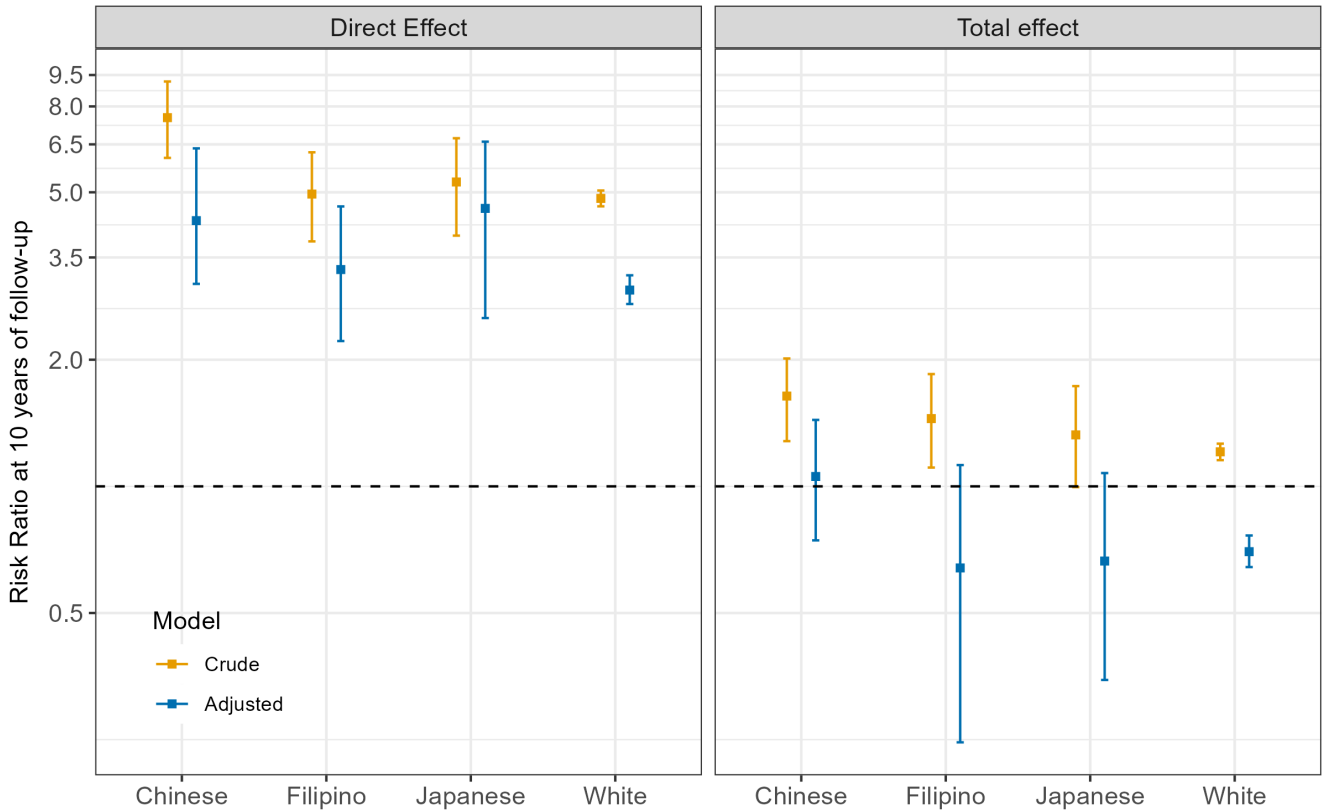
Assumption	Total Effect	Controlled direct effect
Exchangeability	Not needed	Death is independent of future outcomes had everyone followed A = a and death was eliminated, conditional on covariates
Positivity	Not needed	At every follow-up time, there are individuals with any possibly observed level A = a and covariate history who remain alive and free of dementia diagnosis.
Consistency	Not needed	An intervention that “eliminates death” is well-defined.

Smoking cessation on dementia risk over 20 years



Rojas-Saunero et al. AJE.2023

Incident stroke on dementia risk in Asian American and White population in California



Rojas-Saunero et al. *Work in progress.*

Why this matters in health disparities research

Why this matters in health disparities research

- Mortality rates are often higher for marginalized populations.

Why this matters in health disparities research

- Mortality rates are often higher for marginalized populations.
- Descriptive, predictive or causal estimation of dementia risk will be impacted by differential mortality.

Why this matters in health disparities research

- Mortality rates are often higher for marginalized populations.
- Descriptive, predictive or causal estimation of dementia risk will be impacted by differential mortality.
- We can either allow the disparity effect measure to be impacted by the effect of death or we can try to imagine an scenario were we could have eliminated death.

The controlled direct effect estimand

This author views the result in much the same light as a discussion of survival after the zombie apocalypse. The second problem is that the computation for this hypothetical case is only correct

Survival R Package documentation. Thernau 2023.

The controlled direct effect estimand

This author views the result in much the same light as a discussion of survival after the zombie apocalypse. The second problem is that the computation for this hypothetical case is only correct

Survival R Package documentation. Thernau 2023.

In replacing dead participants by cloning the living, IPW generates a sample in which participants are not allowed to die. Moreover, IPW attributes particularly high weights to the participants most likely to die, ie, to people with poor health characteristics associated with death in the attrition model. In doing so, IPW not only prevents people from dying but also artificially maintains the lives of people in very poor health—arguably a form of statistical cruelty.

Chaix et al. Epidemiology.2012

The controlled direct effect estimand

This author views the result in much the same light as a discussion of survival after the zombie apocalypse. The second problem is that the computation for this hypothetical case is only correct

Survival R Package documentation. Thernau 2023.

In replacing dead participants by cloning the living, IPW generates a sample in which participants are not allowed to die. Moreover, IPW attributes particularly high weights to the participants most likely to die, ie, to people with poor health characteristics associated with death in the attrition model. In doing so, IPW not only prevents people from dying but also artificially maintains the lives of people in very poor health—arguably a form of statistical cruelty.

Chaix et al. Epidemiology.2012

1. Do not condition on the future;
2. Do not regard individuals at risk after they have died; and
3. Stick to this world.

Andersen & Keiding. Statistics in Medicine. 2012.

History of competing events analysis

- During the smallpox epidemic (18th century), smallpox inoculation was a controversial therapy.

History of competing events analysis

- During the smallpox epidemic (18th century), smallpox inoculation was a controversial therapy.
- Bernoulli compared the observed life expectancy to a counterfactual scenario that **eliminates smallpox deaths** assuming a hypothethical scenario were every person was innoculated at birth.

History of competing events analysis

- During the smallpox epidemic (18th century), smallpox inoculation was a controversial therapy.
- Bernoulli compared the observed life expectancy to a counterfactual scenario that **eliminates smallpox deaths** assuming a hypothethical scenario were every person was innoculated at birth.
- Therefore smallpox inoculation (as a vaccine) would improve life expectancy for the population at that time.

Karn. 1931; Colombo & Diamanti. 2015

Other possible estimands

- **Composite outcome of dementia and death**

Other possible estimands

- **Composite outcome of dementia and death**
- **Survivors average causal effect:**

$$Pr[Y_t^{a=1} = 1 | \bar{D}_t^{a=1} = \bar{D}_t^{a=0} = 0] - Pr[Y_t^{a=0} = 1 | \bar{D}_t^{a=1} = \bar{D}_t^{a=0} = 0]$$

Other possible estimands

- **Composite outcome of dementia and death**
- **Survivors average causal effect:**

$$Pr[Y_t^{a=1} = 1 | \bar{D}_t^{a=1} = \bar{D}_t^{a=0} = 0] - Pr[Y_t^{a=0} = 1 | \bar{D}_t^{a=1} = \bar{D}_t^{a=0} = 0]$$

- **Natural direct effects**:

$$Pr[Y_t^{a=1, \bar{D}_t^{a=0}} = 1] - Pr[Y_t^{a=0, \bar{D}_t^{a=0}} = 1]$$

Other possible estimands

- **Composite outcome of dementia and death**
- **Survivors average causal effect:**

$$Pr[Y_t^{a=1} = 1 | \bar{D}_t^{a=1} = \bar{D}_t^{a=0} = 0] - Pr[Y_t^{a=0} = 1 | \bar{D}_t^{a=1} = \bar{D}_t^{a=0} = 0]$$

- **Natural direct effects**:

$$Pr[Y_t^{a=1,\bar{D}_t^{a=0}} = 1] - Pr[Y_t^{a=0,\bar{D}_t^{a=0}} = 1]$$

- **Separable effects**
  - Physical decomposition of the exposure assumed to operate on dementia and death through separate pathways. (*Stensrud et al. JASA. 2020*)

Conclusion

- When competing events are present there is not one but several way to consider them as part of the primary research question.

Conclusion

- When competing events are present there is not one but several way to consider them as part of the primary research question.
- Being explicit about the question, allows us to understand the potential sources of bias, and helps us choose between estimators.

Conclusion

- When competing events are present there is not one but several way to consider them as part of the primary research question.
- Being explicit about the question, allows us to understand the potential sources of bias, and helps us choose between estimators.
- We can extend these concepts to exploring the selective process before inclusion into the study.

To search for all the refutable consequences of a hypothesis demands highly imaginative thinking. Imagination is needed to arrive at the hypothesis in the first place, let alone to suggest rigorous tests for it. (Carol Buck)

Buck C. Popper's Philosophy for Epidemiologists, IJE. 1975

Thank you! Gracias!

lp.rojassaunero@ucla.edu

@palolili23

@palolili23