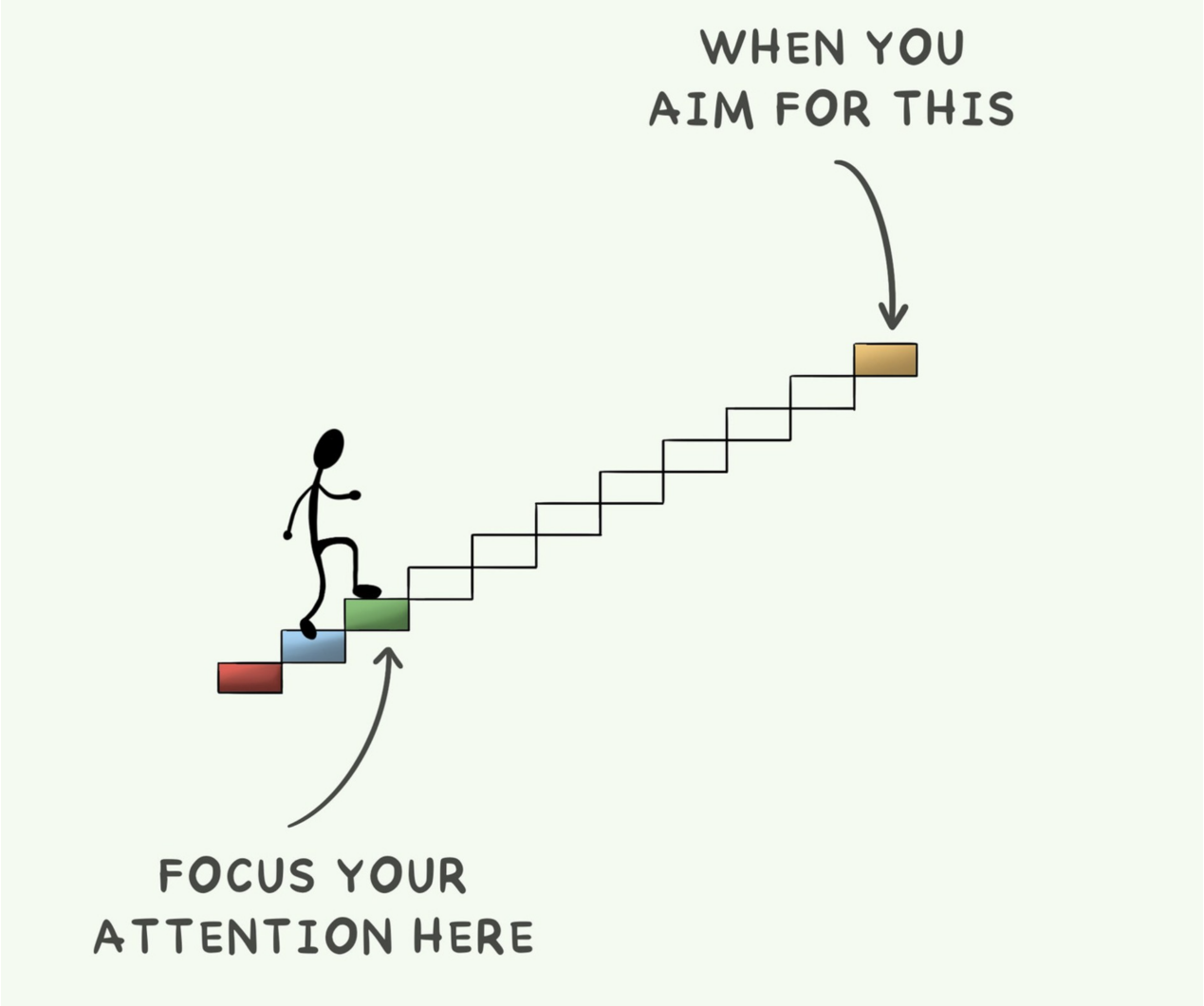
Foundations and conceptual rationale for causal mediation analysis

Mats Stensrud

École Polytechnique Fédérale de Lausanne (EPFL)

Department of Mathematics

Basel 2025



Let's not start estimating something before thinking carefully about the question, estimand and assumptions

The effect of A on Y

Evaluate the effect of some exposure A on outcome Y.

Easy to identify when A is randomly assigned.

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Intensive Lipid Lowering with Atorvastatin in Patients with Stable Coronary Disease

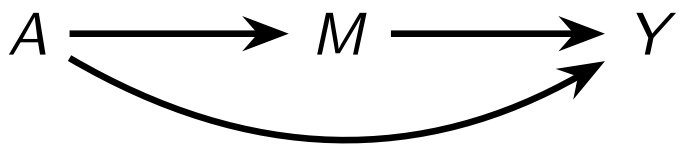
John C. LaRosa, M.D., Scott M. Grundy, M.D., Ph.D.,
David D. Waters, M.D., Charles Shear, Ph.D., Philip Barter, M.D., Ph.D.,
Jean-Charles Fruchart, Pharm.D., Ph.D., Antonio M. Gotto, M.D., D.Phil.,
Heiner Greten, M.D., John J.P. Kastelein, M.D., James Shepherd, M.D.,
and Nanette K. Wenger, M.D., for the Treating to New Targets (TNT) Investigators*

ABSTRACT

Mechanistic effects (Mediation)

Suppose we have established that A affects Y.

Then, many investigators ask about the effect of A on Y,
through or outside of an intermediate variable M.



`Spoiler: this is an ambitious task.`

Statins

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Statin Therapy, LDL Cholesterol, C-Reactive Protein, and Coronary Artery Disease

Steven E. Nissen, M.D., E. Murat Tuzcu, M.D., Paul Schoenhagen, M.D., Tim Crowe, B.S., William J. Sasiela, Ph.D., John Tsai, M.D., John Orazem, Ph.D., Raymond D. Magorien, M.D., Charles O'Shaughnessy, M.D., and Peter Ganz, M.D., for the Reversal of Atherosclerosis with Aggressive Lipid Lowering (REVERSAL) Investigators*

ABSTRACT

How much of statin's effect (A) on Coronary Heart Disease (Y) operates through LDL reduction (M)?

Management of depression for people with cancer (SMaRT oncology 1): a randomised trial

Vanessa Strong, Rachel Waters, Carina Hibberd, Gordon Murray, Lucy Wall, Jane Walker, Gillian McHugh, Andrew Walker, Michael Sharpe

Summary

Background Major depressive disorder severely impairs the quality of life of patients with medical disorders such as cancer, but evidence to guide its management is scarce. We aimed to assess the efficacy and cost of a nurse-delivered complex intervention that was designed to treat major depressive disorder in patients who have cancer.

Methods We did a randomised trial in a regional cancer centre in Scotland, UK. 200 outpatients who had cancer with a prognosis of greater than 6 months and major depressive disorder (identified by screening) were eligible and agreed to take part. Their mean age was 56·6 (SD 11·9) years, and 141 (71%) were women. We randomly assigned 99 of these participants to usual care, and 101 to usual care plus the intervention, with minimisation for sex, age, diagnosis, and extent of disease. The intervention was delivered by a cancer nurse at the centre over an average of seven sessions. The primary outcome was the difference in mean score on the self-reported Symptom Checklist-20 depression scale (range 0 to 4) at 3 months after randomisation. Analysis was by intention to treat. This trial is registered as ISRCTN84767225.

Findings Primary outcome data were missing for four patients. For 196 patients for whom we had data at 3 months, the adjusted difference in mean Symptom Checklist-20 depression score, between those who received the intervention and those who did not, was 0·34 (95% CI 0·13–0·55). This treatment effect was sustained at 6 and 12 months. The intervention also improved anxiety and fatigue but not pain or physical functioning. It cost an additional £5278 (US$10 556) per quality-adjusted life-year gained.

Lancet

Smart trial continues

"those in the cognitive behavioural therapy arm were more likely to use antidepressants during follow-up. This led to questions concerning whether the cognitive behavioural therapy intervention had a beneficial effect on depressive symptoms simply because it led to greater antidepressant use, or whether the intervention affected depressive symptoms through other pathways, e.g. by changing the thought and behavioural patterns of the participants."

VanderWeele, 2016

Vaccine effects on post-infection outcomes

One concern about COVID-19 vaccines is the antibody-dependent enhancement (ADE) phenomenon that vaccine could make the subsequent SARS-CoV-2 infection more severe.[26,27] The ADE phenomenon has been reported in studies of Middle East respiratory syndrome-CoV and SARS-CoV vaccines in animal challenge models.[26,27] However, this was not observed in the preclinical study in an immunization-challenge model of rhesus macaques using the same vaccines in the current study or in reports from preclinical studies of other COVID-19 vaccine candidates, including 2 other inactivated COVID-19 vaccines.[10,11,28] Studies have shown that previous infection of SARS-CoV-2 could protect against rechallenge in rhesus macaques.[29,30]

Xia et al, 2020, JAMA, and several analogous settings.

JAMA and COVID

One concern about COVID-19 vaccines is the antibody-dependent enhancement (ADE) phenomenon that vaccine could make the subsequent SARS-CoV-2 infection more severe.[26,27] The ADE phenomenon has been reported in studies of Middle East respiratory syndrome–CoV and SARS-CoV vaccines in animal challenge models.[26,27] However, this was not observed in the preclinical study in an immunization-challenge model of rhesus macaques using the same vaccines in the current study or in reports from preclinical studies of other COVID-19 vaccine candidates, including 2 other inactivated COVID-19 vaccines.[10,11,28] Studies have shown that previous infection of SARS-CoV-2 could protect against rechallenge in rhesus macaques.[29,30]

Another concern with whole-inactivated virus, particularly with alum adjuvant that can induce T helper 2 cell-biased responses, was vaccine-associated enhanced respiratory disease (VAERD). VAERD was reported in young children in the 1960s when whole-inactivated virus vaccine with alum adjuvant was tested for measles and respiratory syncytial virus.[31,32] However, most of the inactivated vaccines against COVID-19 under development used alum adjuvant,[10,11] and no evidence of VAERD has been seen. Instead, alum may reduce immunopathology compared with unadjuvanted coronavirus vaccines.[33] In the current study, notable changes in the lymphocyte subset distribution or various cytokines (including T helper 2 cell-related cytokines IL-4, IL-5, and IL-10) in various vaccine groups or alum-only group were not observed. However, T-cell–mediated immune responses on stimulation were not measured in the current study. Furthermore, alum is the most widely tested adjuvant component and has been commonly used in many types of vaccines on the market.[34] Nevertheless, safety, including the potential possibility of ADE and VAERD, will be closely monitored in the extended follow-up visits as well as in the phase 3 trial.

Xia et al, JAMA, 2020

Causal inference and decision making

Analyses are **practically useful** if they correspond to **practically relevant questions**.

$\Longrightarrow$ We should be precise about and justify the estimand*, that is, the formalization of our research question.

Let's be explicit about the potential outcomes

- Let superscripts denote potential outcomes:

  Y^a is the outcome had A been set to $a \in \{0, 1\}$.
  possibly contrary to fact

  $Y^{a,m}$ is the outcome had A been set to $a \in \{0, 1\}$ and M been set to $m \in \mathcal{M}$.
  possibly contrary to fact

Mediation concerns sequential interventions

Setup

- First set the treatment A to some value.
- Then intervene on the mediator M.
- Finally observe the outcome Y.

We require well defined interventions on both A and M (Consistency, SUTVA).

Specifying the intervention on M

The estimands we will consider are special cases of:

- **Static regime:** set $M = m_0$ for everyone.
- **Dynamic regime:** set $M = d(A, X)$ using past information (e.g., baseline covariates X and treatment A).
- **Stochastic regime:** modify the distribution of M (for example, a shift or resampling rule conditional on (A, X)). But I doubt that this is useful

Controlled

Controlled direct effects

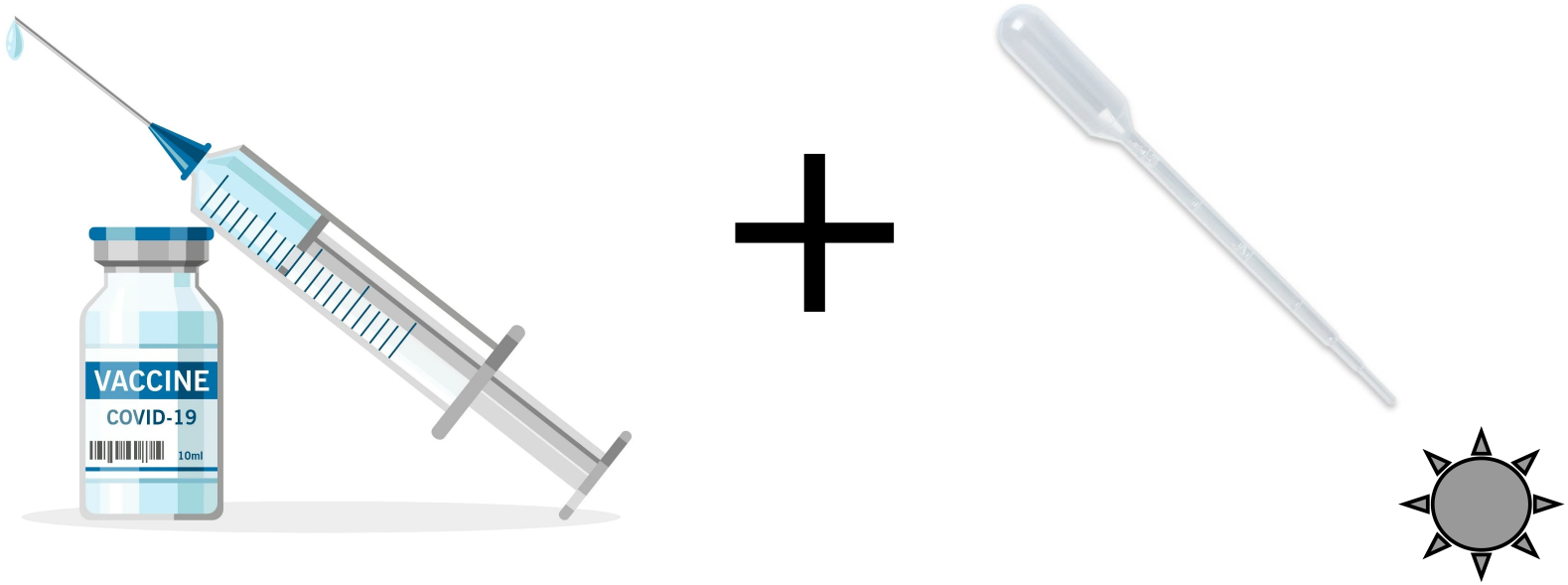
Controlled direct effect at m

$$\mathbb{E}\left(Y^{a=1,\,m}\right) \text{ vs. } \mathbb{E}\left(Y^{a=0,\,m}\right)$$

Interpretation: compare treatment (a=1) to control (a=0) while *holding the mediator fixed at m for everyone*.

How would you define *controlled indirect effects*?

Example 1 of controlled direct effect



$$\mathbb{E}\left(Y^{a=1,\, m=1}\right) \text{ vs. } \mathbb{E}\left(Y^{a=0,\, m=1}\right)$$

Interpretation: compare treatment (a=1) to control (a=0) while *exposing everyone to the infectious agent* ($m = 1$).

Example 1 of controlled direct effect

$$\mathbb{E}\left(Y^{a=1,\, m=1}\right) \text{ vs. } \mathbb{E}\left(Y^{a=0,\, m=1}\right)$$

Interpretation: compare treatment (a=1) to control (a=0) while *exposing everyone to the infectious agent* ($m = 1$).

Example 2 of controlled direct effect

- Take the AIDS Clinical Trials Group 002 trial
  - High- versus low-dose zidovudine
- Administration of prophylaxis therapy for Pneumocystis pneumonia, an opportunistic infection, was left to the physician's discretion. Intercurrent event
- ITT analysis suggested a survival benefit of low-dose zidovudine
- Individuals in the low-dose group received more prophylaxis therapy than those in the high-dose group (61% vs. 50%)
- Would the low-dose group have had better survival than the high-dose group had all trial participants received prophylaxis?
- This question is not addressed by an ITT analysis...

Hernan et al, Annals of internal Medicine, 2013

Natural

Judea Pearl

Strongly advocates the natural direct and indirect effects

Natural direct and indirect effects

Natural direct effect (NDE)

$$\mathbb{E}\left(Y^{a=1,\, M^{a=1}}\right) \quad \text{vs.} \quad \mathbb{E}\left(Y^{a=0,\, M^{a=1}}\right)$$

Interpretation: compare a=1 to a=0 while holding the mediator at the level it would take under a=1.

Natural indirect effect (NIE)

$$\mathbb{E}\left(Y^{a=0,\, M^{a=1}}\right) \quad \text{vs.} \quad \mathbb{E}\left(Y^{a=0,\, M^{a=0}}\right)$$

Interpretation: with A fixed at a=0, compare setting the mediator to the level it would have under a=1 versus a=0.

Note. Even in principle, we cannot design an experiment that directly identifies $\mathbb{E}\left(Y^{a=0,\, M^{a=1}}\right)$, since it combines two incompatible ("cross-world") settings. That concerns me...

Let's take one step back

Formal causal estimands are practically useful if they correspond to practically relevant questions.

Pearl's motivation for natural effects

- Smoking cessation A
- Blood pressure M
- Myocardial infarction Y

Pearl's argued that natural effects have public health importance.

Let's scrutinize that claim

Pearl's modified treatment story

- Pearl's argument:
  - Suppose a new process can completely remove the nicotine from tobacco, allowing the production of a nicotine-free cigarette.
  - Goal: use already collected data on smoking status A, hypertensive status M and MI status Y from a randomized smoking cessation trial to estimate the incidence of MI in smokers were all smokers to change to nicotine-free cigarette.
  - The point is that Pearl tells a story about the **effect of a manipulation that makes no reference to M at all**.

This error was noted by Robins and Richardson.[1]

[1] James M Robins and Thomas S Richardson. "Alternative graphical causal models and the identification of direct effects". In: *Causality and psychopathology: Finding the determinants of disorders and their cures* (2010), pp. 103–158.

Pearl's story about a 4 arm trial

The estimand referred to natural direct and indirect effects of smoking on MI outcomes relative to mediation by blood pressure.

- To defend this effect as "actionable" or "clinically relevant" he told a story about effects of modified cigarettes on MI
- These could be tested in a 4 arm trial where individuals are randomized to
  1. Not smoke
  2. Smoke current cigarettes
  3. Smoke cigarettes with only nicotine removed
  4. Smoke cigarettes with only tobacco removed (e-cigarettes...)

So let's rather target effects of a 4-arm trial

Statins and cardiovascular disease

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Statin Therapy, LDL Cholesterol, C-Reactive Protein, and Coronary Artery Disease

Steven E. Nissen, M.D., E. Murat Tuzcu, M.D., Paul Schoenhagen, M.D., Tim Crowe, B.S., William J. Sasiela, Ph.D., John Tsai, M.D., John Orazem, Ph.D., Raymond D. Magorien, M.D., Charles O'Shaughnessy, M.D., and Peter Ganz, M.D., for the Reversal of Atherosclerosis with Aggressive Lipid Lowering (REVERSAL) Investigators*

ABSTRACT

Would it make sense to create a new drug that only targets LDL?

Yes, PCSK9 inhibitors

Or CRP?

No, I don't think so

Smart trial

Management of depression for people with cancer (SMaRT oncology 1): a randomised trial

Vanessa Strong, Rachel Waters, Carina Hibberd, Gordon Murray, Lucy Wall, Jane Walker, Gillian McHugh, Andrew Walker, Michael Sharpe

Summary

Background Major depressive disorder severely impairs the quality of life of patients with medical disorders such as cancer, but evidence to guide its management is scarce. We aimed to assess the efficacy and cost of a nurse-delivered complex intervention that was designed to treat major depressive disorder in patients who have cancer.

Methods We did a randomised trial in a regional cancer centre in Scotland, UK. 200 outpatients who had cancer with a prognosis of greater than 6 months and major depressive disorder (identified by screening) were eligible and agreed to take part. Their mean age was 56·6 (SD 11·9) years, and 141 (71%) were women. We randomly assigned 99 of these participants to usual care, and 101 to usual care plus the intervention, with minimisation for sex, age, diagnosis, and extent of disease. The intervention was delivered by a cancer nurse at the centre over an average of seven sessions. The primary outcome was the difference in mean score on the self-reported Symptom Checklist-20 depression scale (range 0 to 4) at 3 months after randomisation. Analysis was by intention to treat. This trial is registered as ISRCTN84767225.

Findings Primary outcome data were missing for four patients. For 196 patients for whom we had data at 3 months, the adjusted difference in mean Symptom Checklist-20 depression score, between those who received the intervention and those who did not, was 0·34 (95% CI 0·13–0·55). This treatment effect was sustained at 6 and 12 months. The intervention also improved anxiety and fatigue but not pain or physical functioning. It cost an additional £5278 (US$10 556) per quality-adjusted life-year gained.

Lancet

JAMA and COVID

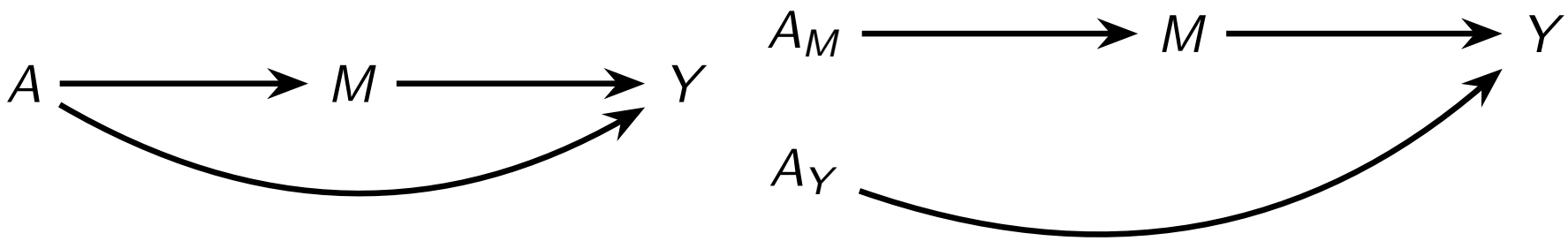
One concern about COVID-19 vaccines is the antibody-dependent enhancement (ADE) phenomenon that vaccine could make the subsequent SARS-CoV-2 infection more severe.[26,27] The ADE phenomenon has been reported in studies of Middle East respiratory syndrome–CoV and SARS-CoV vaccines in animal challenge models.[26,27] However, this was not observed in the preclinical study in an immunization-challenge model of rhesus macaques using the same vaccines in the current study or in reports from preclinical studies of other COVID-19 vaccine candidates, including 2 other inactivated COVID-19 vaccines.[10,11,28] Studies have shown that previous infection of SARS-CoV-2 could protect against rechallenge in rhesus macaques.[29,30]

Another concern with whole-inactivated virus, particularly with alum adjuvant that can induce T helper 2 cell-biased responses, was vaccine-associated enhanced respiratory disease (VAERD). VAERD was reported in young children in the 1960s when whole-inactivated virus vaccine with alum adjuvant was tested for measles and respiratory syncytial virus.[31,32] However, most of the inactivated vaccines against COVID-19 under development used alum adjuvant,[10,11] and no evidence of VAERD has been seen. Instead, alum may reduce immunopathology compared with unadjuvanted coronavirus vaccines.[33] In the current study, notable changes in the lymphocyte subset distribution or various cytokines (including T helper 2 cell-related cytokines IL-4, IL-5, and IL-10) in various vaccine groups or alum-only group were not observed. However, T-cell–mediated immune responses on stimulation were not measured in the current study. Furthermore, alum is the most widely tested adjuvant component and has been commonly used in many types of vaccines on the market.[34] Nevertheless, safety, including the potential possibility of ADE and VAERD, will be closely monitored in the extended follow-up visits as well as in the phase 3 trial.

Xia et al, JAMA, 2020

Separable

Let's be explicit



Hypothesize components of A, that is: A_M and A_Y

Separable effects

Separable direct effect

$$\mathbb{E}\left(Y^{a_Y=1,\, a_M=1}\right) \quad \text{vs.} \quad \mathbb{E}\left(Y^{a_Y=0,\, a_M=1}\right)$$

Effect of the A_Y component of A.

Effect of A on Y not through M.

Separable indirect effect

$$\mathbb{E}\left(Y^{a_Y=0,\, a_M=1}\right) \quad \text{vs.} \quad \mathbb{E}\left(Y^{a_Y=0,\, a_M=0}\right)$$

Effect of the A_M component of A.

Effect of A on Y only through M.

Inspired by Robins & Richardson (2010), Didelez (2020). See also Stensrud et al., JASA (2020).

We have considered simplifications

Treatments (like, Statins) are often taken sequentially $(A_0, A_1, ...)$.
The so-called mediator is often a process too $(M_0, M_1, ...)$.

- We need to take the time-varying structure seriously
- I have not even talked about identification

  Natural effects generally require cross-world assumptions that are **untestable even in principle**.

Some important points

- Choose and justify estimands on a case-by-case basis
  - The estimands should reflect real-life questions
  - Helps us to avoid logical flaws and erroneous thinking
  - Sharpens arguments and thought processes
- Useful estimands share some key features
  - Defined in observable subsets of the population
  - Can be empirically falsified in a future experiment

Exercise for the upcoming talks. Ask:

- Are the interventions well defined?
- What is the policy relevance of the estimand?
- Can we rephrase the question in practical terms?
- Are the assumptions credible?

Some references

1. Stensrud MJ et al. *Conditional separable effects* Journal of the American Statistical Association (2022)
2. Sarvet AL, Stensrud MJ. *Without commitment to an ontology there could be no causal inference*. Epidemiology (2022).
3. Stensrud MJ et al. *Separable Effects for Causal Inference in the Presence of Competing Risks*. Journal of the American Statistical Association (2020).
4. Stensrud MJ et al. *Generalized interpretation and identification of separable effects in competing risk settings* Lifetime data analysis and https://arxiv.org/abs/2004.14824 (2020)
5. Janvin M, Young JG, Ryalen PC, Stensrud MJ. *Causal inference with recurrent and competing events*. arXiv preprint arXiv:2202.08500. (2022)
6. Stensrud MJ, Dukes O. *Translating questions to estimands in randomized clinical trials with intercurrent events*. Statistics in Medicine (2022).
7. Stensrud MJ, Young JG, Martinussen T. *Discussion on "Causal mediation of semicompeting risks" by Yen-Tsung Huang*. Biometrics. 2021 Dec;77(4):1160-4.
8. Robins and Richardson. *Alternative graphical causal models and the identification of direct effects*. Causality and psychopathology: Finding the determinants of disorders and their cures (2010): 103-158.
9. Didelez V *Defining causal mediation with a longitudinal mediator and a survival outcome*. Lifetime data analysis (2018)
10. Robins JM et al. *An interventionist approach to mediation analysis* https://arxiv.org/abs/2008.06019

Appendix

ACS and outcomes

SECTION 1: AN ILLUSTRATIVE CASE

Go to:

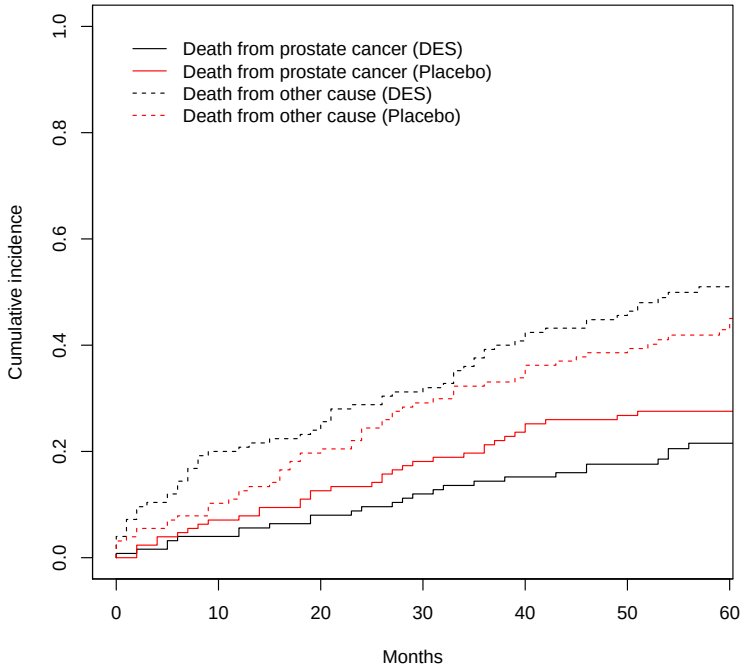
Acute coronary syndrome (ACS) presents as a cardiac emergency caused by sudden obstruction of a coronary artery, most frequently due to thrombus formation in an existing atherosclerotic lesion in the vessel wall. In the acute phase, treatment aims to prevent sudden cardiac death and complications by halting the progression of thrombus formation, managing symptoms, and identifying and treating coronary obstructions; the latter goal involves early cardiac catheterization. Once stabilized, patients receive secondary preventive medication and undergo risk factor modification to prevent future cardiovascular events, including death.

Using Danish register data, we have previously established [3]that in a population of patients with a first hospitalization for ACS, the use of an early invasive treatment strategy was associated with a lower short-term risk of cardiac death and readmission for myocardial infarction than a conservative approach. It has been speculated that some or, in selected subgroups, all of the long-term benefit provided by an invasive treatment strategy is mediated through better secondary preventive medical therapy. In this case study, we will explore the relationship between an early invasive treatment strategy, secondary preventive medication, and death from all causes.

Lange et al, Epidemiology and Health, 2017

A drug to reduce prostate cancer mortality

`Byar and Green (1980)`



- Diethylstilbestrol (DES, *black*) vs Placebo (*red*).
- When more people die from other causes, fewer people can die from prostate cancer.

See Stensrud et al, JASA, 2020