Identification of vaccine effects in time-to-event settings

Mats J. Stensrud

École Polytechnique Fédérale de Lausanne (EPFL)

Department of Mathematics

2021

Joint work with Louisa H. Smith

[Submitted on 22 Nov 2021]

Identification of vaccine effects when exposure status is unknown

Mats J. Stensrud, Louisa H. Smith

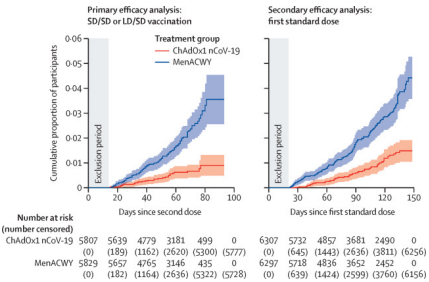
Results from randomized controlled trials (RCTs) help determine vaccination strategies and related public health policies. However, defining and identifying estimands that can guide policies in infectious disease settings is difficult, even in an RCT. The effects of vaccination critically depend on characteristics of the population of interest, such as the prevalence of infection, the number of vaccinated, and social behaviors. To mitigate the dependence on such characteristics, estimands (and study designs) that require conditioning or intervening on exposure to the infectious agent have been advocated. But a fundamental problem for both RCTs and observational studies is that exposure status is often unavailable or difficult to measure, which has made it impossible to apply existing methodology to study vaccine effects that account for exposure status. In this work, we present new results on this type of vaccine effects. Under plausible conditions, we show that point identification of certain relative effects is possible even when the exposure status is unknown. Furthermore, we derive sharp bounds on the corresponding absolute effects. We apply these results to estimate the effects of the ChAdOx1 nCoV-19 vaccine on SARS-CoV-2 disease (COVID-19) conditional on post-vaccine exposure to the virus, using data from a large RCT.

Mats J. Stensrud and Louisa H. Smith. *Identification of vaccine effects when exposure status is unknown*. 2021. arXiv: 2111.11548 [stat.ME]

Outline

Motivation

Vaccines are one of the biggest successes in medicine



Merryn Voysey et al. "Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK". In: *The Lancet* 397.10269 (2021), pp. 99–111

Policy-relevant estimands in infectious disease settings

...are difficult to define and identify , even with data from an RCT.

They depend on characteristics of the population of interest , such as

- the prevalence of infection,
- the number of vaccinated,
- and social behaviors.

Interference...

How about conditioning, or intervening, on exposure?

Perspective

Accelerating Development of SARS-CoV-2 Vaccines — The Role for Controlled Human Infection Models

Meagan E. Deming, M.D., Ph.D., Nelson L. Michael, M.D., Ph.D., Merlin Robb, M.D., Myron S. Cohen, M.D., and Kathleen M. Neuzil, M.D., M.P.H.


Vaccine 39 (2021) 633–640

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

WHO Report

Key criteria for the ethical acceptability of COVID-19 human challenge studies: Report of a WHO Working Group

Euzebiusz Jamrozik [a,b,c], Katherine Littler [d], Susan Bull [a], Claudia Emerson [e], Gagandeep Kang [f], Melissa Kapulu [g,h], Elena Rey [e,i], Carla Saenz [j], Seema Shah [k], Peter G Smith [l], Ross Upshur [m], Charles Weijer [n], Michael J Selgelid [b], for the WHO Working Group for Guidance on Human Challenge Studies in COVID-19

How about conditioning, or intervening, on exposure when we don't have access to challenge trials?

- Such estimands mitigate dependence on
  - the prevalence of infection,
  - the number of vaccinated,
  - and social behaviors.
- But a fundamental problem: exposure status is often unmeasured.
  - For example, measuring susceptibility to infection "might not be easy in practice and might indeed require considerable assumptions regarding who is infectious and when, how infectious the persons are, and who is exposing whom." halloran1995causalhalloran2010design
    Similarly, identification "generally requires measurement of each person's exposure history during the follow-up, which is rarely, if ever, available." o2014estimating
- What can we infer when the exposure status is **unmeasured**?

How about conditioning, or intervening, on exposure when we don't have access to challenge trials?

- Such estimands mitigate dependence on
  - the prevalence of infection,
  - the number of vaccinated,
  - and social behaviors.
- But a fundamental problem: exposure status is often unmeasured.
  - For example, measuring susceptibility to infection "might not be easy in practice and might indeed require considerable assumptions regarding who is infectious and when, how infectious the persons are, and who is exposing whom." halloran1995causalhalloran2010design
    Similarly, identification "generally requires measurement of each person's exposure history during the follow-up, which is rarely, if ever, available." o2014estimating
- What can we infer when the exposure status is **unmeasured**?

How about conditioning, or intervening, on exposure when we don't have access to challenge trials?

- Such estimands mitigate dependence on
  - the prevalence of infection,
  - the number of vaccinated,
  - and social behaviors.
- But a fundamental problem: exposure status is often unmeasured.
  - For example, measuring susceptibility to infection "might not be easy in practice and might indeed require considerable assumptions regarding who is infectious and when, how infectious the persons are, and who is exposing whom." halloran1995causalhalloran2010design
    Similarly, identification "generally requires measurement of each person's exposure history during the follow-up, which is rarely, if ever, available." o2014estimating
- What can we infer when the exposure status is **unmeasured**?

How about conditioning, or intervening, on exposure when we don't have access to challenge trials?

- Such estimands mitigate dependence on
  - the prevalence of infection,
  - the number of vaccinated,
  - and social behaviors.
- But a fundamental problem: exposure status is often unmeasured.
  - For example, measuring susceptibility to infection "might not be easy in practice and might indeed require considerable assumptions regarding who is infectious and when, how infectious the persons are, and who is exposing whom." halloran1995causalhalloran2010design
    Similarly, identification "generally requires measurement of each person's exposure history during the follow-up, which is rarely, if ever, available." o2014estimating
- What can we infer when the exposure status is **unmeasured**?

How about conditioning, or intervening, on exposure when we don't have access to challenge trials?

- Such estimands mitigate dependence on
  - the prevalence of infection,
  - the number of vaccinated,
  - and social behaviors.
- But a fundamental problem: exposure status is often unmeasured.
  - For example, measuring susceptibility to infection
    "might not be easy in practice and might indeed require considerable assumptions regarding who is infectious and when, how infectious the persons are, and who is exposing whom."[1][2]
    Similarly, identification "generally requires measurement of each person's exposure history during the follow-up, which is rarely, if ever, available."[3]
- What can we infer when the exposure status is **unmeasured**?

[1] M Elizabeth Halloran and Claudio J Struchiner. "Causal inference in infectious diseases". In: *Epidemiology* (1995), pp. 142–151.

[2] M Elizabeth Halloran, Claudio J Struchiner, and Ira M Longini Jr. *Design and analysis of vaccine studies*. Vol. 18. Springer, 2010.

[3] Justin J O'Hagan, Marc Lipsitch, and Miguel A Hernán. "Estimating the per-exposure effect of infectious disease interventions". In: *Epidemiology (Cambridge, Mass.)* 25.1 (2014), p. 134.

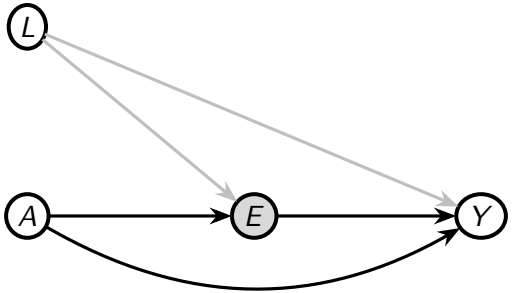
Let's be explicit

Consider a blinded RCT where individuals are drawn from a much larger population , so that interactions among patients in the trial are negligible , where

- $A \in \{0, 1\}$ is a vaccine (treatment) indicator.
- $L \in \mathcal{L}$ is a vector of baseline covariates.
- $E \in \{0, 1\}$ is an exposure (to virus) indicator, which is unmeasured.
- $Y \in \{0, 1, \ldots, M\}$ is the event of interest (severity of infection).

Let superscripts denote counterfactuals , such that

- E^a is exposure to COVID-19 had, possibly contrary to fact, A been set to $a \in \{0, 1\}$.
- Y^a is severe COVID-19 infection status had, possibly contrary to fact, A been set to $a \in \{0, 1\}$.



(a)

Estimands with different interpretations

- The average treatment effect (ATE),

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

- A naive contrast of counterfactual outcomes conditional on exposure status,

$$\mathbb{E}(Y^{a=1} \mid E^{a=1} = 1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E^{a=0} = 1).$$

- The principal stratum effect (PSE),

$$\mathbb{E}(Y^{a=1} \mid E^{a=0} = E^{a=1} = 1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E^{a=0} = E^{a=1} = 1).$$

- The causal effect conditional on observed exposure,

$$\mathbb{E}(Y^{a=1} \mid E = 1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E = 1).$$

- The controlled direct effect (CDE),

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

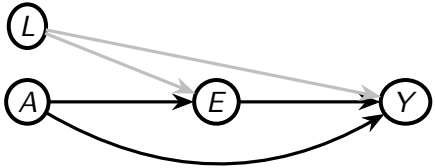
A setting where exposure status is unaffected by treatment

A *blinded* RCT, which is the context of many vaccine efficacy studies.

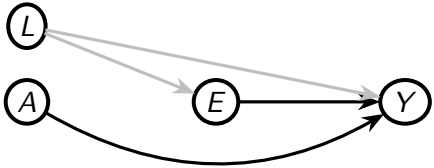
Assumption (No effect on exposure)

$$E^{a=0} = E^{a=1}.$$

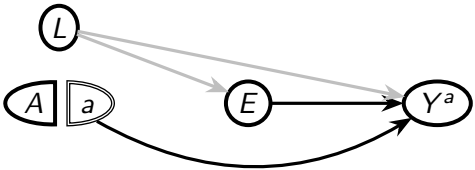
Causal graphs encoding the assumptions



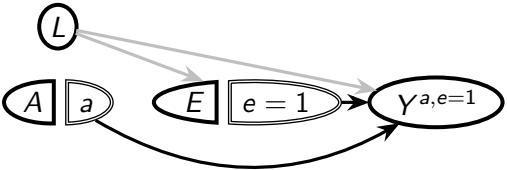
(b)



(c)



(d)



(e)

Equalities under the "no effect on exposure assumptions"

When $E^{a=0} = E^{a=1}$,

$$\begin{aligned} &\mathbb{E}(Y^{a=1} \mid E = 1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E = 1) \\ = &\mathbb{E}(Y^{a=1} \mid E^{a=1} = 1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E^{a=0} = 1) \\ = &\mathbb{E}(Y^{a=1} \mid E^{a=0} = E^{a=1} = 1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E^{a=0} = E^{a=1} = 1) \end{aligned}$$

I will denote these contrasts the **Causal Effect Conditional on Exposure (CECE).**

Identification

The ATE is easy to identify

Assumption (Treatment exchangeability)

$$Y^a, E^a \perp\!\!\!\perp A$$

Assumption (Positivity)

$$P(A = a) > 0 \quad \forall a \in \{0, 1\}$$

Assumption (Consistency)

$$\textit{If } A = a, \textit{ then } E = E^a, Y = Y^a.$$

These conditions allow us to identify the ATE as $\mathbb{E}(Y \mid A = 1)$ vs. $\mathbb{E}(Y \mid A = 0)$, regardless of whether exposure status E is measured.

Consider the Causal Effect Conditional on Exposure (CECE)

It is straightforward to express the CECE as a functional of *factual* variables,

$$\begin{aligned} &\mathbb{E}(Y^{a=1} \mid E=1) \text{ vs. } \mathbb{E}(Y^{a=0} \mid E=1) \\ =\ &\mathbb{E}(Y \mid E=1, A=1) \text{ vs. } \mathbb{E}(Y \mid E=1, A=0), \end{aligned}$$

using the assumptions of Exchangeability, Positivity and Consistency.

But this functional is *not identified* in our data because $\mathbb{E}(Y \mid E=1, A=a)$ is not estimable when E is unmeasured.

To be clear, consider the additive CECE

$$\mathbb{E}(Y^{a=1} \mid E=1) - \mathbb{E}(Y^{a=0} \mid E=1) = \mathbb{E}(Y \mid E=1, A=1) - \mathbb{E}(Y \mid E=1, A=0).$$

Exposure to the virus is necessary for having severe infection

Assumption (Exposure necessity)

$$E^a = 0 \implies Y^a = 0, \ \forall a \in \{0, 1\}.$$

Let's use "No effect on exposure" and "Exposure necessity"

Theorem (Relative CECE)

Under standard identifiability conditions for ATEs, exposure necessity and the no effect on exposure assumption, the relative CECE is equal to

$$\frac{\mathbb{E}(Y^{a=1} \mid E=1)}{\mathbb{E}(Y^{a=0} \mid E=1)} = \frac{\mathbb{E}(Y \mid A=1)}{\mathbb{E}(Y \mid A=0)}.$$

Theorem (Absolute CECE)

Under the same assumptions, the absolute CECE is partially identified by the sharp bounds

$$\mathbb{E}(Y \mid A=0) - \mathbb{E}(Y \mid A=1) \leq \mathbb{E}(Y^{a=0} \mid E=1) - \mathbb{E}(Y^{a=1} \mid E=1) \leq 1 - \frac{\mathbb{E}(Y \mid A=1)}{\mathbb{E}(Y \mid A=0)}.$$

Relation to common definitions of vaccine effects

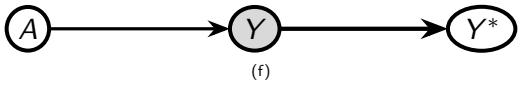
Level	Parameter Choice	Comparison Groups and Effect: Susceptibility	Infectiousness	Combined Change in Susceptibility and Infectiousness
	Conditional on exposure to infection:			
I	Transmission probability, p Secondary attack rate, SAR	$\mathrm{VE}_{S,p}\dagger = 1 - \frac{p_{10}}{p_{11}}$	$\mathrm{VE}_{I,p} = 1 - \frac{p_{11}}{p_{01}}$	$\mathrm{VE}_{T,p} = 1 - \frac{p_{11}}{p_{00}}$

		Study Design: I Direct	IIa Indirect	IIb Total	III Overall
	Unconditional:				
II	Incidence rate, IR	$\mathrm{VE}_{S,IR} = 1 - \frac{\mathrm{IR}_{A1}}{\mathrm{IR}_{A0}}$	$\mathrm{VE}_{IIa,IR} = 1 - \frac{\mathrm{IR}_{A0}}{\mathrm{IR}_{B0}}$	$\mathrm{VE}_{IIb,IR} = 1 - \frac{\mathrm{IR}_{A1}}{\mathrm{IR}_{B0}}$	$\mathrm{VE}_{III,IR} = 1 - \frac{\mathrm{IR}_{A\cdot}}{\mathrm{IR}_{B\cdot}}$
	Hazard rate, λ	$\mathrm{VE}_{S,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{A0}}$	$\mathrm{VE}_{IIa,\lambda} = 1 - \frac{\lambda_{A0}}{\lambda_{B0}}$	$\mathrm{VE}_{IIb,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{B0}}$	$\mathrm{VE}_{III,\lambda} = 1 - \frac{\lambda_{A\cdot}}{\lambda_{B\cdot}}$
III	Proportional hazards, PH	$\mathrm{VE}_{S,PH} = 1 - e^{\beta_1}$	NA	NA	NA
IV	Cumulative incidence, CI Attack rates, AR	$\mathrm{VE}_{S,CI} = 1 - \frac{\mathrm{CI}_{A1}}{\mathrm{CI}_{A0}}$	$\mathrm{VE}_{IIa,CI} = 1 - \frac{\mathrm{CI}_{A0}}{\mathrm{CI}_{B0}}$	$\mathrm{VE}_{IIb,CI} = 1 - \frac{\mathrm{CI}_{A1}}{\mathrm{CI}_{B0}}$	$\mathrm{VE}_{III,CI} = 1 - \frac{\mathrm{CI}_{A\cdot}}{\mathrm{CI}_{B\cdot}}$

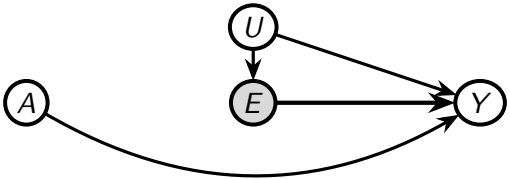
M Elizabeth Halloran, Claudio J Struchiner, and Ira M Longini Jr. *Design and analysis of vaccine studies*. Vol. 18. Springer, 2010

By the way, this is related to a fun fact in epidemiology

- Diagnostic tests that have perfect specificity give unbiased estimates of risk ratios, even if these tests mis-classify disease cases (we need that $A \perp\!\!\!\perp Y^* \mid Y$).



(f)

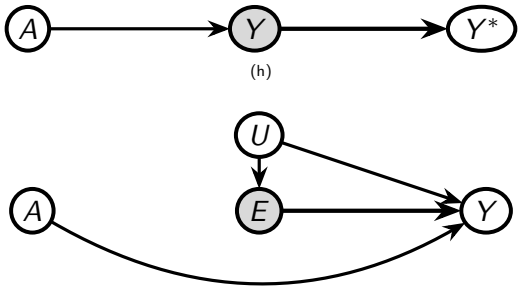


(g)

- See also Zhao et al zhao2020note who studied racial discrimination in policing.

By the way, this is related to a fun fact in epidemiology

- Diagnostic tests that have perfect specificity give unbiased estimates of risk ratios, even if these tests mis-classify disease cases (we need that $A \perp\!\!\!\perp Y^* \mid Y$).

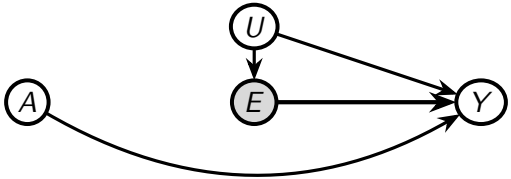


- See also Zhao et al zhao2020note who studied racial discrimination in policing.

By the way, this is related to a fun fact in epidemiology

- Diagnostic tests that have perfect specificity give unbiased estimates of risk ratios, even if these tests mis-classify disease cases (we need that $A \perp\!\!\!\perp Y^* \mid Y$).

(j)



(k)

- See also Zhao et al[4] who studied racial discrimination in policing.

[4] Qingyuan Zhao et al. "A note on post-treatment selection in studying racial discrimination in policing". In: *arXiv preprint arXiv:2009.04832* (2020).

So, how about the controlled direct effect (CDE)?

Assumption (Exposure exchangeability)

$$Y^{a,e=1} \perp\!\!\!\perp A \mid L \quad \text{and} \quad Y^{a,e=1} \perp\!\!\!\perp E^a \mid L, A.$$

Assumption (Exposure positivity)

$$P(A = a, E = 1 \mid L) > 0 \quad \forall a \in \{0, 1\} \ \textit{w.p.1.}$$

Assumption (Exposure consistency)

$$\textit{If } A = a \textit{ and } E = 1 \textit{ then } Y = Y^{a,e=1}.$$

Under these conditions the CDE can be expressed as

$$\mathbb{E}(Y^{a,e=1}) = \mathbb{E}\{\mathbb{E}(Y \mid E = 1, A = a, L)\},$$

which is well-known, but here we cannot identify $\mathbb{E}(Y \mid E = 1, A = a, L)$ ...

The CDE conditional on L can also be identified

Theorem (CDE conditional on L)

Under classical conditions for CDE, exposure necessity and the no effect on exposure assumption, the relative CDE conditional on the baseline covariate L is

$$\frac{\mathbb{E}(Y^{a=1,e=1} \mid L)}{\mathbb{E}(Y^{a=0,e=1} \mid L)} = \frac{\mathbb{E}(Y \mid A=1, L)}{\mathbb{E}(Y \mid A=0, L)}.$$

Corollary (CECE and CDE conditional on L)

Under the same conditions, the relative CECE given $L = l$ and the relative CDE conditional on the baseline covariate $L = l$ are equal, that is,

$$\frac{\mathbb{E}(Y^{a=1} \mid E^{a=1}=1, L=l)}{\mathbb{E}(Y^{a=0} \mid E^{a=0}=1, L=l)} = \frac{\mathbb{E}(Y^{a=1,e=1} \mid L=l)}{\mathbb{E}(Y^{a=0,e=1} \mid L=l)} = \frac{\mathbb{E}(Y \mid A=1, L=l)}{\mathbb{E}(Y \mid A=0, L=l)}.$$

Point identification of the absolute CECE

Suppose we have access to $\mathbb{E}(Y \mid E = 1, A = a)$ or $P(E = 1 \mid A = a)$, e.g.

Corollary (Point identification of the absolute CECE)

Under the same assumptions as for the CECE theorems,

$$\begin{aligned}
&\mathbb{E}(Y^{a=0} \mid E = 1) - \mathbb{E}(Y^{a=1} \mid E = 1) \\
=&\mathbb{E}(Y \mid E = 1, A = 0)\left(1 - \frac{\mathbb{E}(Y \mid A = 1)}{\mathbb{E}(Y \mid A = 0)}\right) && (1) \\
=&\frac{\mathbb{E}(Y \mid A = 0)}{P(E = 1 \mid A = 0)} - \frac{\mathbb{E}(Y \mid A = 1)}{P(E = 1 \mid A = 1)}. && (2)
\end{aligned}$$

Time

Time-to-events

Consider a blinded RCT, where

- $A \in \{0, 1\}$ is a vaccine (treatment) indicator at baseline,
- E_k indicates exposure by time interval $k = 0, 1, 2..., K$ (e.g. to the coronavirus), which is still unmeasured.
- Y_k indicates the outcome by time interval $= 0, 1, 2..., K$ (e.g. severe infection).
- C_k indicate loss to follow-up (censoring) by time $k = 0, 1, 2..., K$.

"No effect on exposure" and "Exposure necessity" in time-to-event settings

Assumption (Time-varying exposure necessity)

$$E_k^{a,\overline{c}=0} = 0 \implies Y_k^{a,\overline{c}=0} = 0.$$

Assumption (No effect on exposure)

$$E_k^{a=0,\overline{c}=0} = E_k^{a=1,\overline{c}=0}.$$

Theorem (Relative and absolute CECE for time to event outcomes)

Under standard ATE conditions, exposure necessity and the no effect on exposure assumption for time-to-event outcomes,

$$\frac{\mathbb{E}(Y_k^{a=1,\overline{c}=0} \mid E_k^{a=1,\overline{c}=0} = 1)}{\mathbb{E}(Y_k^{a=0,\overline{c}=0} \mid E_k^{a=0,\overline{c}=0} = 1)} = \frac{\mu_k(1)}{\mu_k(0)},$$

where

$$\mu_k(a) = \sum_{s=1}^{k} h_s(a) \prod_{j=0}^{s-1} [1 - h_j(a)]$$

and

$$h_k(a) = \frac{\mathbb{E}[Y_k(1 - Y_{k-1})(1 - C_k) \mid A = a]}{\mathbb{E}[(1 - Y_{k-1})(1 - C_k) \mid A = a]}.$$

Under the same conditions, the absolute CECE is partially identified by the sharp bounds

$$\mu_k(0) - \mu_k(1) \leq \mathbb{E}(Y_k^{a=0,\overline{c}=0} \mid E_k^{a=0,\overline{c}=0} = 1) - \mathbb{E}(Y_k^{a=1,\overline{c}=0} \mid E_k^{a=1,\overline{c}=0} = 1) \leq 1 - \frac{\mu_k(1)}{\mu_k(0)}.$$

A Single World Intervention Graph

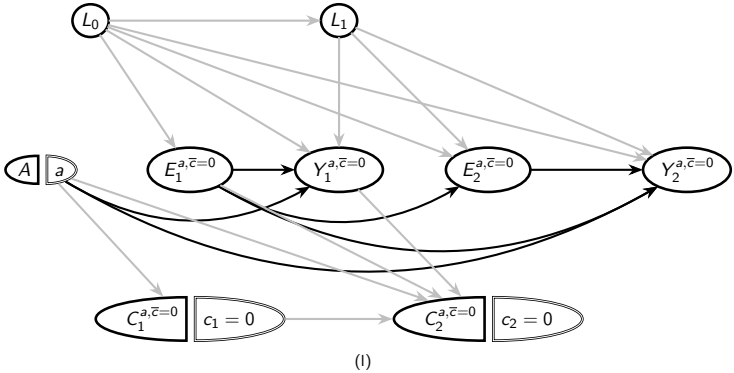


Figure: The SWIG shows a time-to-event setting where the CECE is identified, even if L_0 and L_1 are unmeasured.

Excess fraction vs. etiologic fraction

The excess fraction among the exposed in interval k is

$$\begin{aligned}&\frac{\mathbb{E}(Y_k^{a=0,c=0} \mid E_k^{a=0,c=0}=1) - \mathbb{E}(Y_k^{a=1,c=0} \mid E_k^{a=1,c=0}=1)}{\mathbb{E}(Y_k^{a=0,c=0} \mid E_k^{a=0,c=0}=1)} \\ =&\frac{\mathbb{E}(Y_k \mid A=0) - \mathbb{E}(Y_k \mid A=1)}{\mathbb{E}(Y_k \mid A=0)} = 1 - \frac{\mathbb{E}(Y_k \mid A=1)}{\mathbb{E}(Y_k \mid A=0)},\end{aligned}$$

which quantifies the proportionate increase in caseload under no treatment.
Not to be confused with the etiologic fraction, i.e. the fraction *caused* (or prevented) by treatment, which is equivalent to the probability of causation.[5]

[5] Sander Greenland and James M Robins. "Conceptual problems in the definition and interpretation of attributable fractions.". In: *American journal of epidemiology* 128.6 (1988), pp. 1185–1197.

Estimation

Estimates in the example

- Let $\hat{\mu}(a)$ be an estimator of $\mathbb{E}(Y \mid A = a)$, e.g. an empirical mean.
- We estimate the relative CECE by

$$\widehat{\text{rCECE}} = \frac{\hat{\mu}(1)}{\hat{\mu}(0)},$$

- We estimate bounds of the absolute CECE by
  $\widehat{\text{aCECE}}_U = 1 - \widehat{\text{rCECE}}$.
  $\widehat{\text{aCECE}}_L = \hat{\mu}(0) - \hat{\mu}(1)$.
- Similarly, Let $\hat{\mu}(a, l)$ be an estimator of $\mathbb{E}(Y \mid A = a, L = l)$, and we can derive results conditional on L.

Example

ChAdOx1 nCoV-19 vaccine against COVID-19

- Blinded RCT done across the UK, Brazil, and South Africa.[6]
- Recruited 60-90% health-care workers, depending on the site.
- ChAdOx1 nCoV-19 vaccine or control (which contained a meningococcal vaccine).
- The interim analysis included 11636 participants.
- The cumulative incidence of COVID-19 80 days since second dose was
  - 0.9% (95% CI: $0.5\% - 1.3\%$) in the vaccine arm.
  - 3.1% (95% CI : $2.4\% - 3.8\%$) in the placebo arm.
  - Thus, an estimate of the relative CECE $\equiv$ $\text{CECE}_{k=80}$ is

$$\widehat{\text{rCECE}} = \frac{\hat{\mu}_1}{\hat{\mu}_0} = 0.30 \text{ (95\% CI: } 0.15 - 0.44),$$

corresponding to the reported vaccine efficacy point estimate of $1 - 0.30 = 0.70$.

[6] Merryn Voysey et al. "Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK". In: *The Lancet* 397.10269 (2021), pp. 99–111.

What do the bounds tell us

- The sharp lower bound

$$\widehat{\text{aCECE}}_L = 0.031 - 0.009 = 0.022 \text{ (95\% CI: } 0.011 - 0.033),$$

  which is reached under a setting where everybody is exposed to the virus.
- The sharp upper bound

$$\widehat{\text{aCECE}}_U = 1 - 0.30 = 0.70 \text{ (95\% CI: } 0.57 - 0.85),$$

  which is reached when then probability of the outcome among the exposed is 1. Not very informative...

Can we reason about $P(E = 1 \mid A = 0)$?

- Suppose 60% of the trial participants were exposed to a specific amount of virus particles such that the exposure necessity condition holds. Then, $\widehat{\text{aCECE}} = 0.037$.

  which would imply that $P(Y = 1 \mid E = 1, A = 0) = 0.052$;
  that is E was sufficient to cause symptomatic COVID-19 in just over 5% of unvaccinated participants during 80 days of follow-up.

- Suppose $P(E = 1 \mid A = 0)$ to 0.9 (instead of 0.6).Then, $\widehat{\text{aCECE}} = 0.024$

  which would imply that $P(Y = 1 \mid E = 1, A = 0) = 0.034$.

Can we reason about $P(Y = 1 \mid E = 1, A = 0)$?

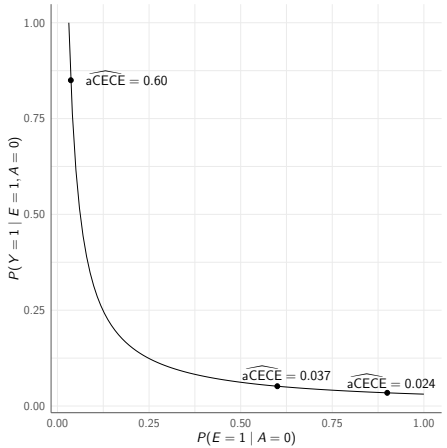
- Consider the choir practice in Washington state in March 2020hamner2020high,
  after which 52 out of 61 participants developed COVID-19,
  having been exposed to a high concentration of virus in an unmasked setting .
- If exposure to such a high dose of SARS-CoV-2 is necessary for infection,
  - $P(Y = 1 \mid E = 1, A = 0) = 0.85$, and
  - $\widehat{\text{aCECE}} = 0.60$,
  - an estimate consistent with $P(E = 1 \mid A = 0) = 0.036$.

Can we reason about $P(Y = 1 \mid E = 1, A = 0)$?

- Consider the choir practice in Washington state in March 2020[7],
  after which 52 out of 61 participants developed COVID-19,
  having been exposed to a high concentration of virus in an unmasked setting .
- If exposure to such a high dose of SARS-CoV-2 is necessary for infection,
  - $P(Y = 1 \mid E = 1, A = 0) = 0.85$, and
  - $\widehat{\text{aCECE}} = 0.60$,
  - an estimate consistent with $P(E = 1 \mid A = 0) = 0.036$.

[7] Lea Hamner. "High SARS-CoV-2 attack rate following exposure at a choir practice—Skagit County, Washington, March 2020". In: *MMWR. Morbidity and mortality weekly report* 69 (2020).

Illustration of bounds in the ChAdOx1 nCoV-19 vaccine



Conclusion

Conclusion

- Publicized results are often *relative effects*, as in major studies on COVID-19 vaccines.
  - Our results give new interpretations to numbers people often compute.
- The results are perhaps surprising.
  - Re: estimating vaccine effects conditional on exposure status "requires information on who is infectious and when, and whom they contact and how" halloran2010design.
- We require that exposure (say, close contact with an infectious individual) is *necessary* for the outcome of interest to occur (say, symptomatic disease).
  - Alternatively, we could adapt the definition of exposure to something measurable.
  - For example, close contact with infected people who present overt disease.
  - But such definitions have explicitly been discouraged, because they would lead to an *underestimate* of the exposure in settings where inapparent infections exist.

Conclusion

- Publicized results are often *relative effects*, as in major studies on COVID-19 vaccines.
  - Our results give new interpretations to numbers people often compute.
- The results are perhaps surprising.
  - Re: estimating vaccine effects conditional on exposure status "requires information on who is infectious and when, and whom they contact and how" halloran2010design.
- We require that exposure (say, close contact with an infectious individual) is *necessary* for the outcome of interest to occur (say, symptomatic disease).
  - Alternatively, we could adapt the definition of exposure to something measurable.
  - For example, close contact with infected people who present overt disease.
  - But such definitions have explicitly been discouraged, because they would lead to an *underestimate* of the exposure in settings where inapparent infections exist.

Conclusion

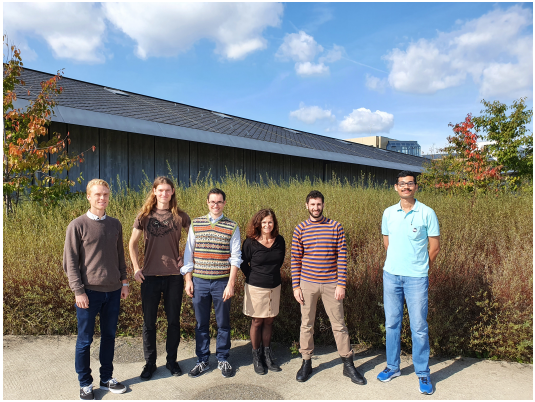
- Publicized results are often *relative effects*, as in major studies on COVID-19 vaccines.
  - Our results give new interpretations to numbers people often compute.
- The results are perhaps surprising.
  - Re: estimating vaccine effects conditional on exposure status "requires information on who is infectious and when, and whom they contact and how"[8].
- We require that exposure (say, close contact with an infectious individual) is *necessary* for the outcome of interest to occur (say, symptomatic disease).
  - Alternatively, we could adapt the definition of exposure to something measurable.
  - For example, close contact with infected people who present overt disease.
  - But such definitions have explicitly been discouraged, because they would lead to an *underestimate* of the exposure in settings where inapparent infections exist.

[8] Halloran, Struchiner, and Longini Jr, *Design and analysis of vaccine studies*.

Future directions

- Formally consider generalizability and transportability.
- Relate these results (and assumptions) to results (and assumptions) in the infectious disease modelling literature.

My group at EPFL (2021)



Please get in touch if you would like to work with us