Causal mediation analysis in molecular biology:

Characterizing immune correlates of vaccine protection

Nima Hejazi

Wednesday, 7th October 2026

Department of Biostatistics
T.H. Chan School of Public Health
Harvard University

nimahejazi.org
nhejazi | nshlab

EMBO course: Causality in Biomedicine
European Bioinformatics Institute (EMBL-EBI)
Joint work with P. Gilbert and M. van der Laan

Identifying immune correlates of vaccination for HIV-1 and COVID-19

The fights against HIV-1 and COVID-19

- The HIV-1 epidemic:
  - 1.5 million new infections occurring annually worldwide;
  - new infections outpace patients starting antiretroviral therapy;
  - HIV Vaccine Trials Network's (HVTN) *505* trial evaluated a novel antibody boost vaccine (Hammer et al. 2013).
- The COVID-19 ~~epi~~ ~~pan~~ endemic (Antia and Halloran 2021):
  - ~~270~~ ~~331~~ ~~619~~ ~~643~~ 770 million total cases detected globally;
  - new variants emerging, with continued formulation of targeted vaccines and annual roll-outs expected;
  - COVID-19 Prevention Network's (CoVPN) *COVE* trial focused on Moderna's (mRNA-1273) vaccine (Baden et al. 2021).

Evaluating vaccines for HIV-1 and COVID-19

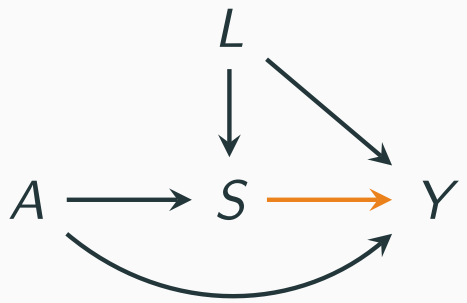
- *505*: How would HIV-1 infection risk have differed had the boost vaccine modulated antibody responses differently?
- *COVE*: How would COVID-19 disease risk have differed for alternative vaccine-induced immunogenic response profiles?
- **Question**: How can [HIV-1, COVID-19] vaccines be improved through the modulation of immunogenic response profiles?

Why measure and analyze immune correlates?

- Two, interrelated goals of immune correlates analyses are to
  - identify/validate possible *surrogate endpoints* (Prentice 1989);
  - understand/delineate *protective mechanisms* of vaccines.
- If an immune correlate is established to reliably predict VE, subsequent efficacy trials may use it as a primary endpoint.
- This may accelerate the approval of
  - existing vaccines in different populations (e.g., in children);
  - new vaccines within the same class (e.g., bivalent mRNA);
  - inform the development of "next-generation" vaccines.

Correlates of risk versus correlates of protection

- Correlate of risk (CoR): an immune marker S *associated* with the endpoint Y among vaccine recipients.
- Correlate of protection (CoP): an immune marker that is *causally* predictive of protection against Y.
- A CoR need not be a CoP: baseline factors L (e.g., age, behavioral risk) may drive both immune response and disease risk.
- Is S a CoP? Question of causal mediation: how much of the effect of A on Y operates through S? (Gilbert et al. 2024)



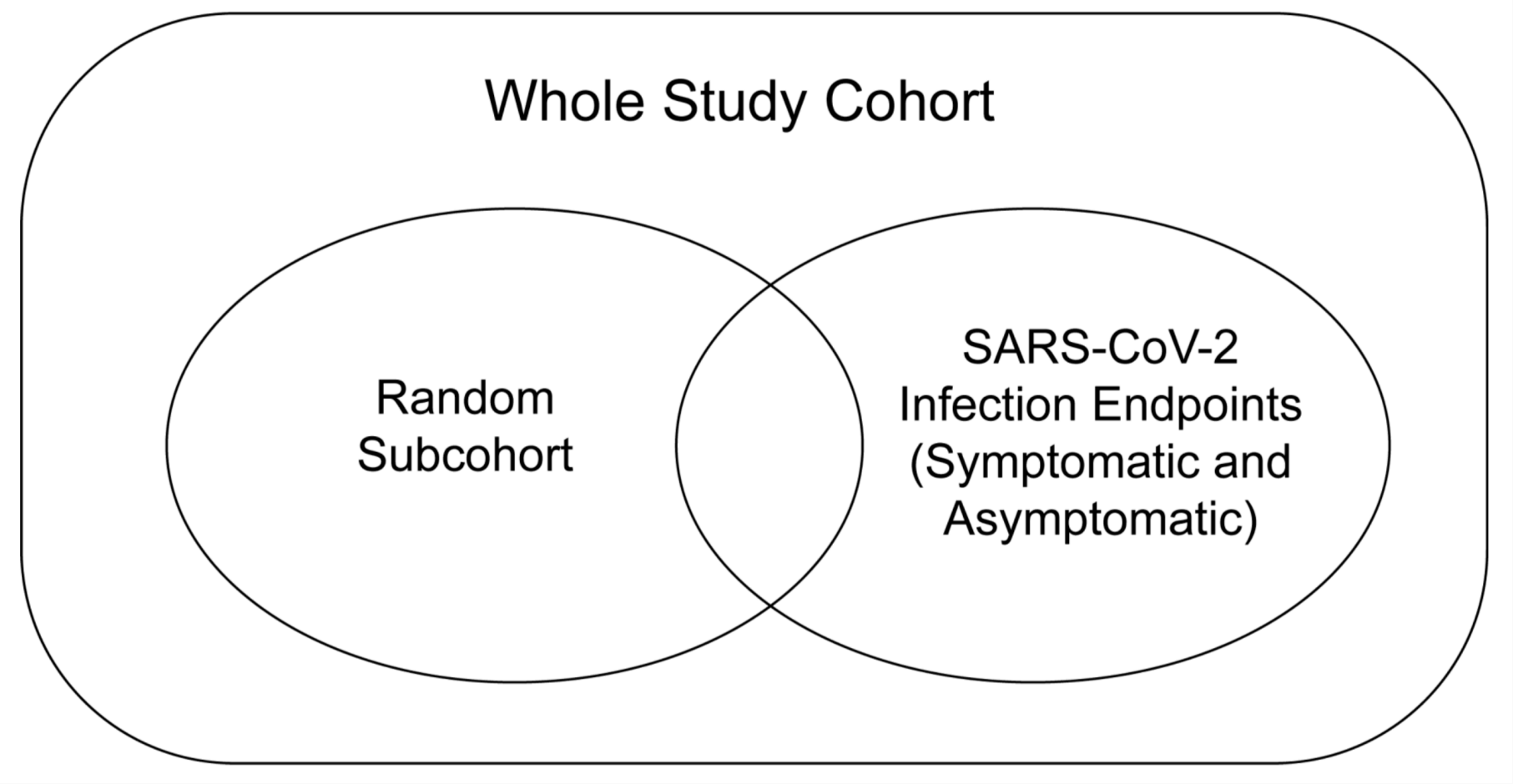
A is randomized, but the $S \rightarrow Y$ path is subject to confounding by L.

Measuring correlates: Two-phase designs

- Often, use case-cohort design (Prentice 1986), a special case of two-phase sampling (Breslow et al. 2003) commonly used in vaccine trials (Follmann 2006).
- Phase 1: measure baseline, vaccination, endpoint on everyone.
- Phase 2: given baseline, vaccine, endpoint, select members of immune response subcohort with (possibly known) probability.
  - *505*: phase-two sample with 100% of HIV-1 cases and matching of non-cases ($n = 189$ per Janes et al. 2017).
  - *COVE*: stratified random subcohort ($n \approx 1600$) and all SARS-CoV-2 infection and COVID-19 disease endpoints.

A simple two-phase design: Case-cohort

Assaying >30k samples is expensive, statistically unnecessary.



Case-cohort design, per Prentice (1986), as applied to COVE.

Two-phase sampling masks the complete data structure

- Complete (unobserved) data $X = (L, A, S, Y) \sim \mathrm{P}_0^X \in \mathcal{P}$:
  - L (baseline covariates): sex, age, BMI, behavioral HIV risk,
  - A (treatment): randomized assignment to vaccine/placebo,
  - S (exposure): immune response profile for relevant markers,
  - Y (outcome of interest): infection status at trial's end.
- Observed data $O = (V, R, RX) = (L, A, RS, Y, R) \sim \mathrm{P}_0 \in \mathcal{P}$.
  - $V \equiv (L, Y)$ are used in defining *outcome-dependent* two-phase sampling mechanism $g_{0,R} := \mathbb{P}(R = 1 \mid V)$.
  - $R \in \{0, 1\}$ is an indicator of inclusion in the phase-two sample.
  - $g_{0,R} := \mathbb{P}(R = 1 \mid Y, L)$ must be *known by design* or estimated.

Causal effects for continuous exposures

Static interventions aren't enough

- Describe the manner in which X is hypothetically generated by a structural causal model (Pearl 2009):

  $$L = f_L(U_L); A \sim \text{Bern}(0.5); S = f_S(A, L, U_S); Y = f_Y(S, A, L, U_Y)$$

- SCM implies a model for the distribution of counterfactual RVs induced by interventions on the system under study.
- A *static* intervention would replace f_S with a specific value s in its conditional support, i.e., $S \mid A = 1, L$.
- This requires specifying *a priori* a particular value under which to evaluate the outcome—but what value?

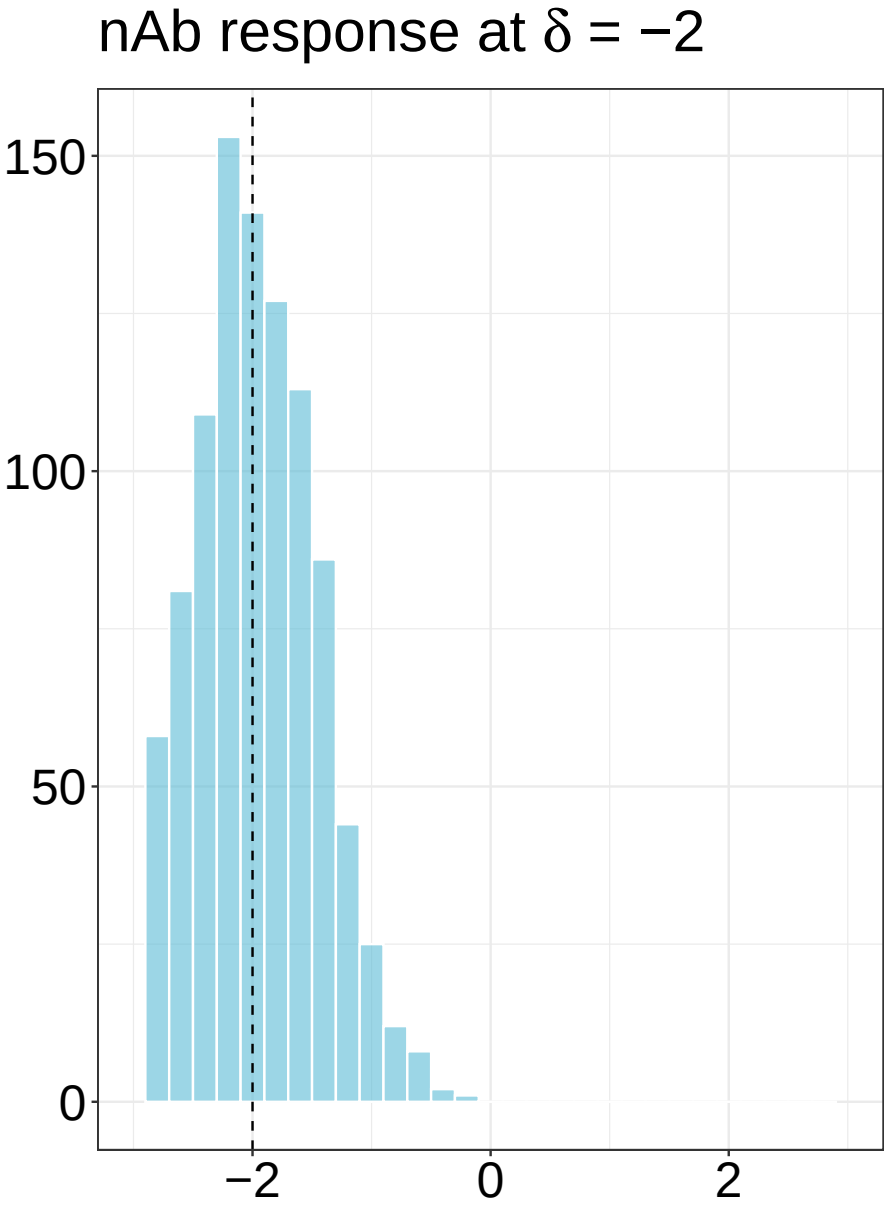
Controlled vaccine efficacy (CVE)

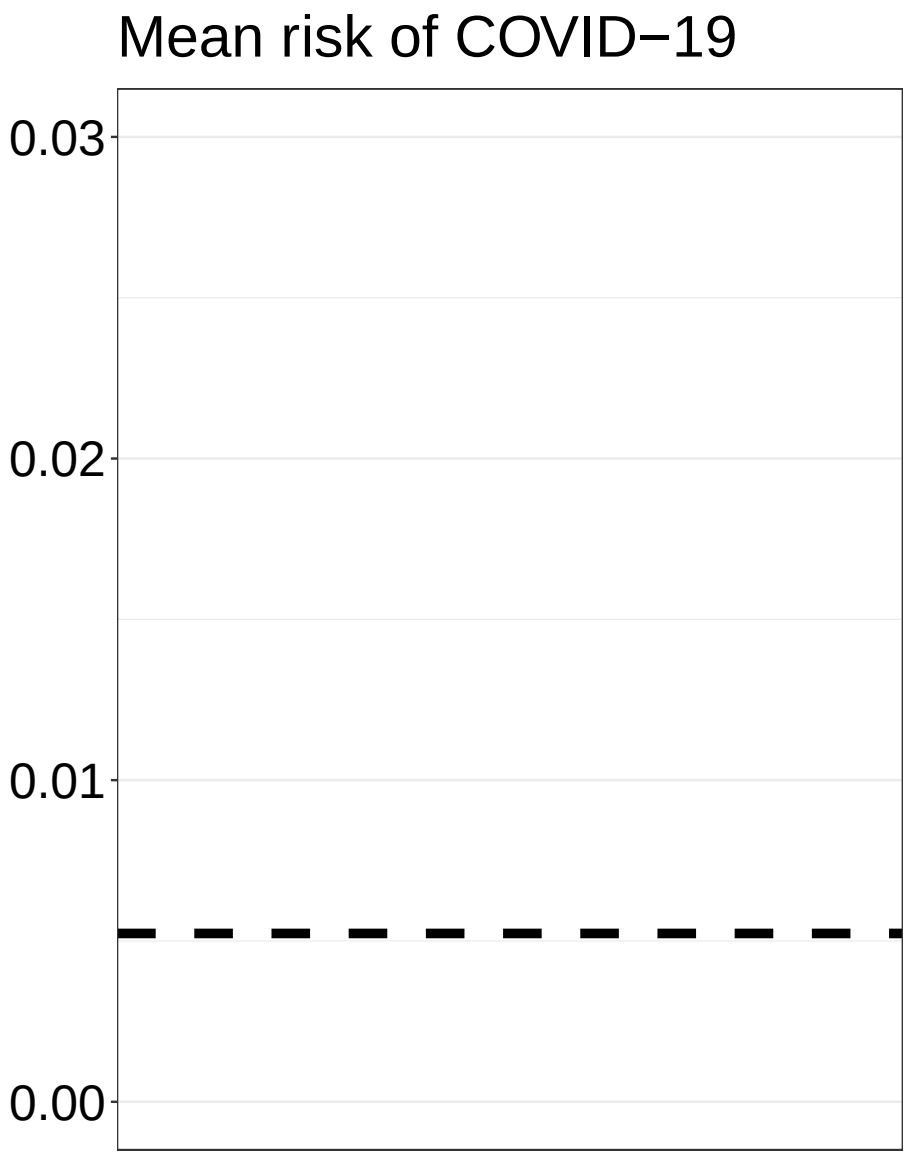
- For a hypothetical value $s \in \mathcal{S}$, the *controlled direct effect* (CDE) quantifies the effect of A on Y while fixing $S = s$.
- The hypothetical value $S = s$ must be chosen carefully—to be scientifically informative *and* avoid positivity violations.
- For two hypothetical values $s_0, s_1 \in \mathcal{S}$, *Controlled Vaccine Efficacy* (CVE) (Gilbert et al. 2023) is defined

$$\text{CVE}(s_0, s_1) = 1 - \frac{\mathbb{E}[\mathbb{P}(Y = 1 \mid S = s_1, A = 1, L)]}{\mathbb{E}[\mathbb{P}(Y = 1 \mid \cancel{S = s_0}, A = 0, L)]},$$

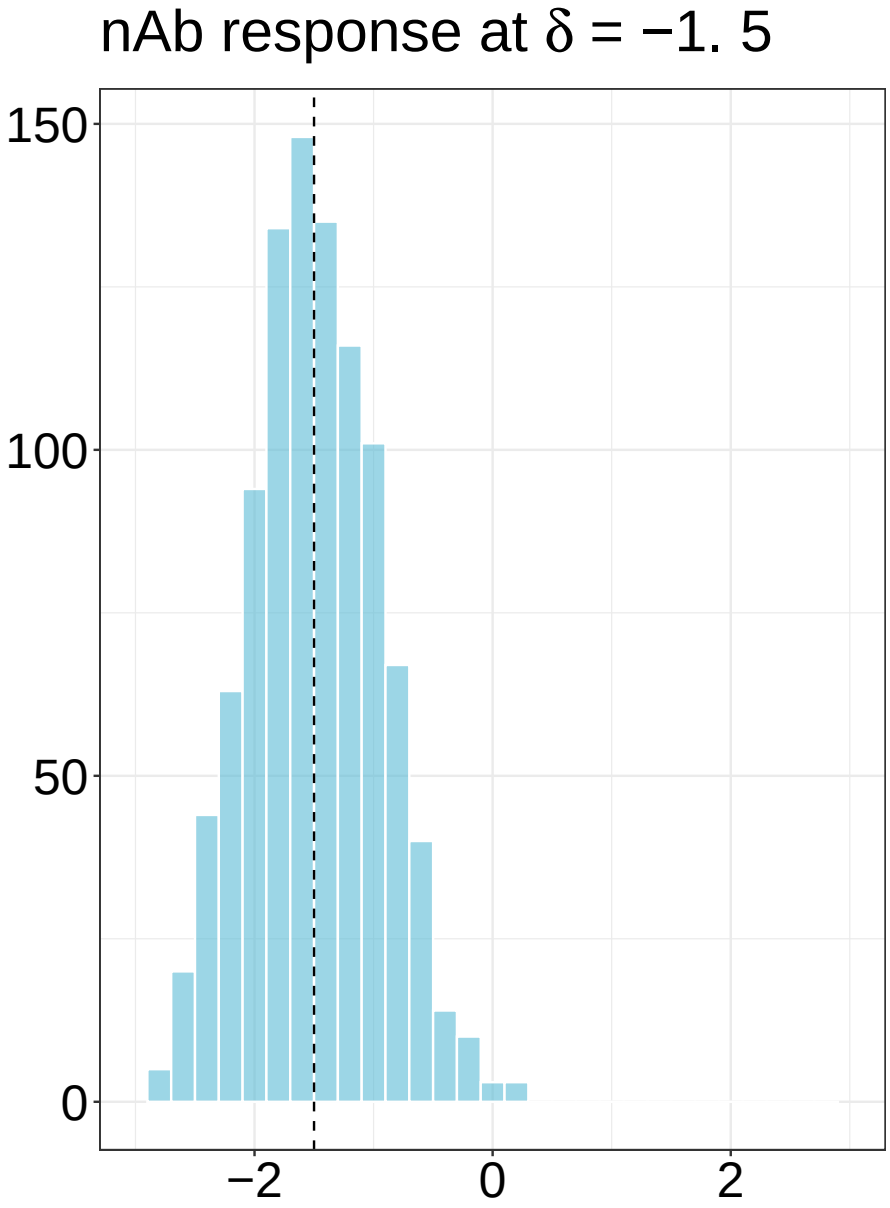
  which contrasts counterfactual risk for vaccine and $S = s_1$ vs. placebo, where $s_0 = 0$ is argued a plausible assumption in pathogen-naive populations (Gilbert et al. 2024).

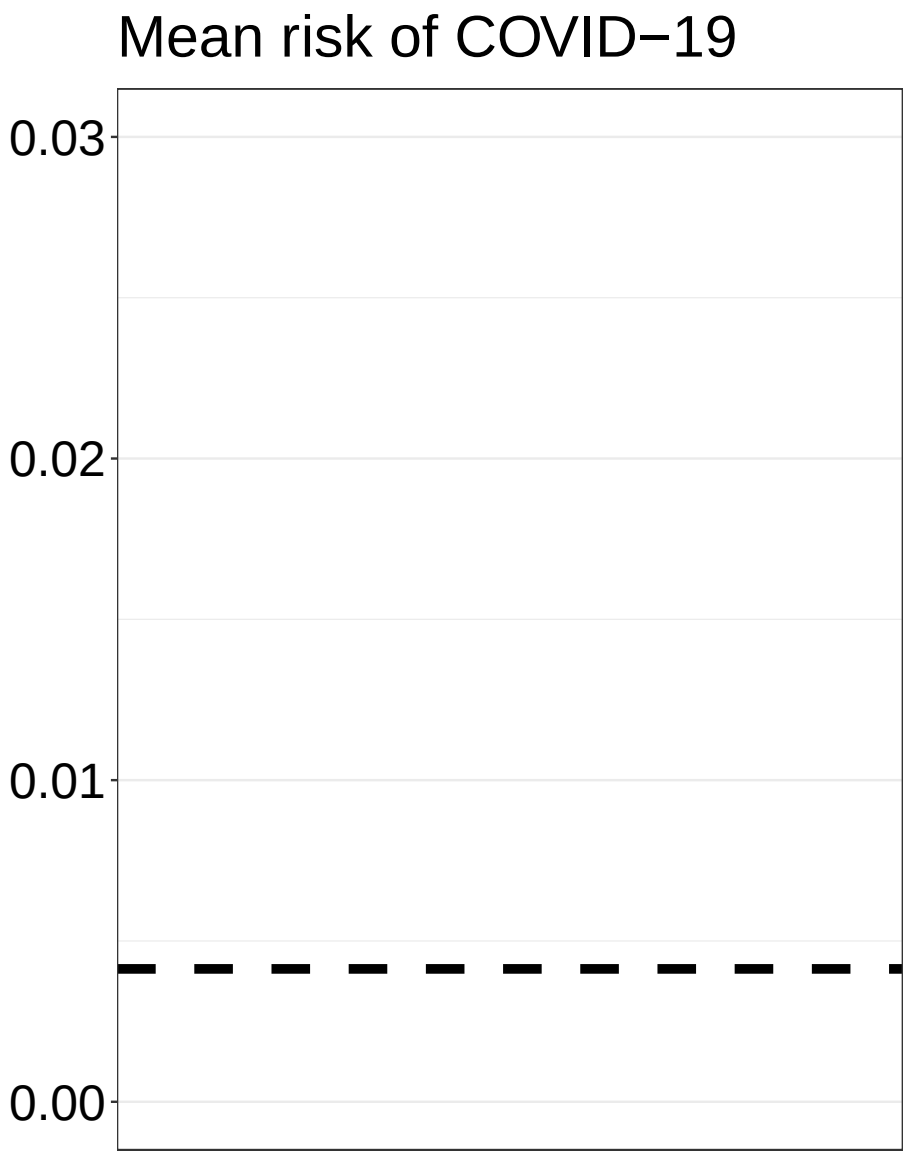
Disease risk under shifted immunogenic responses



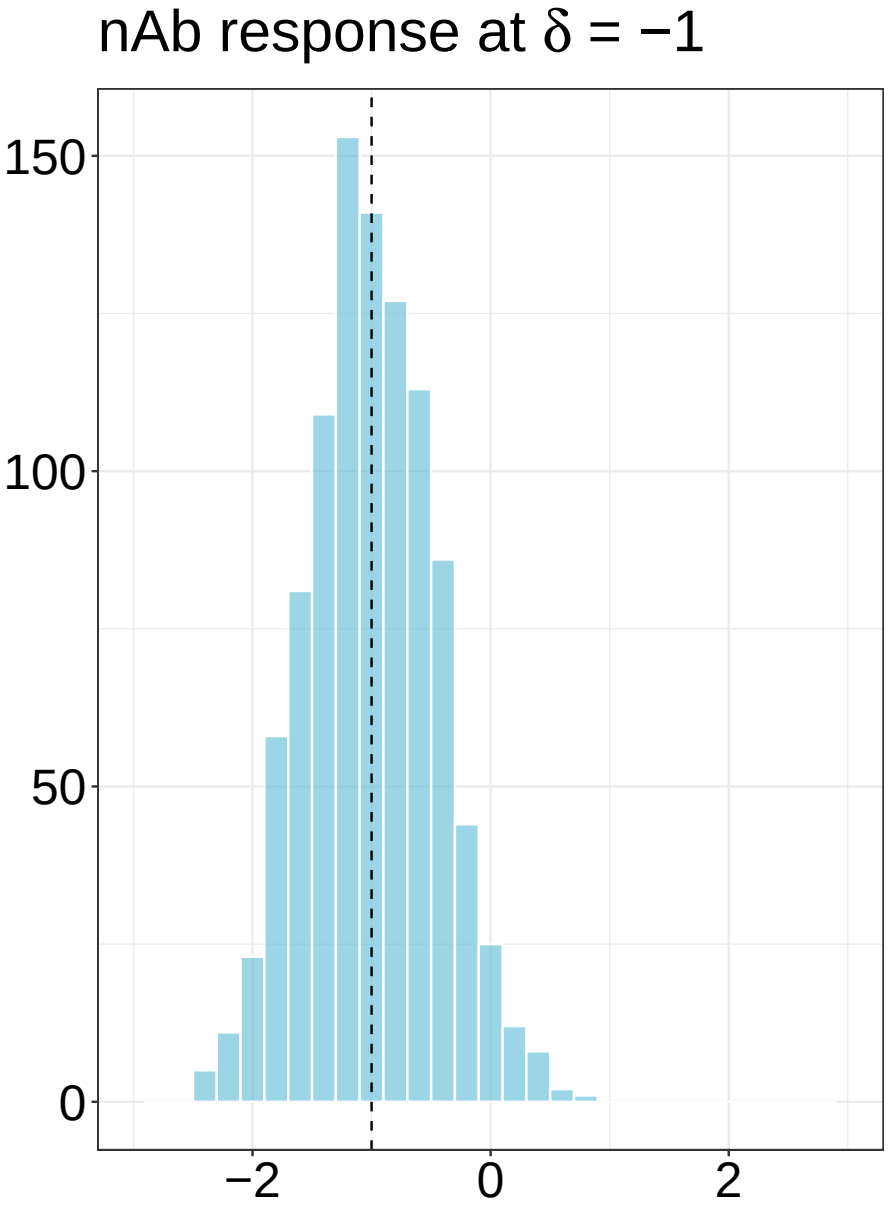


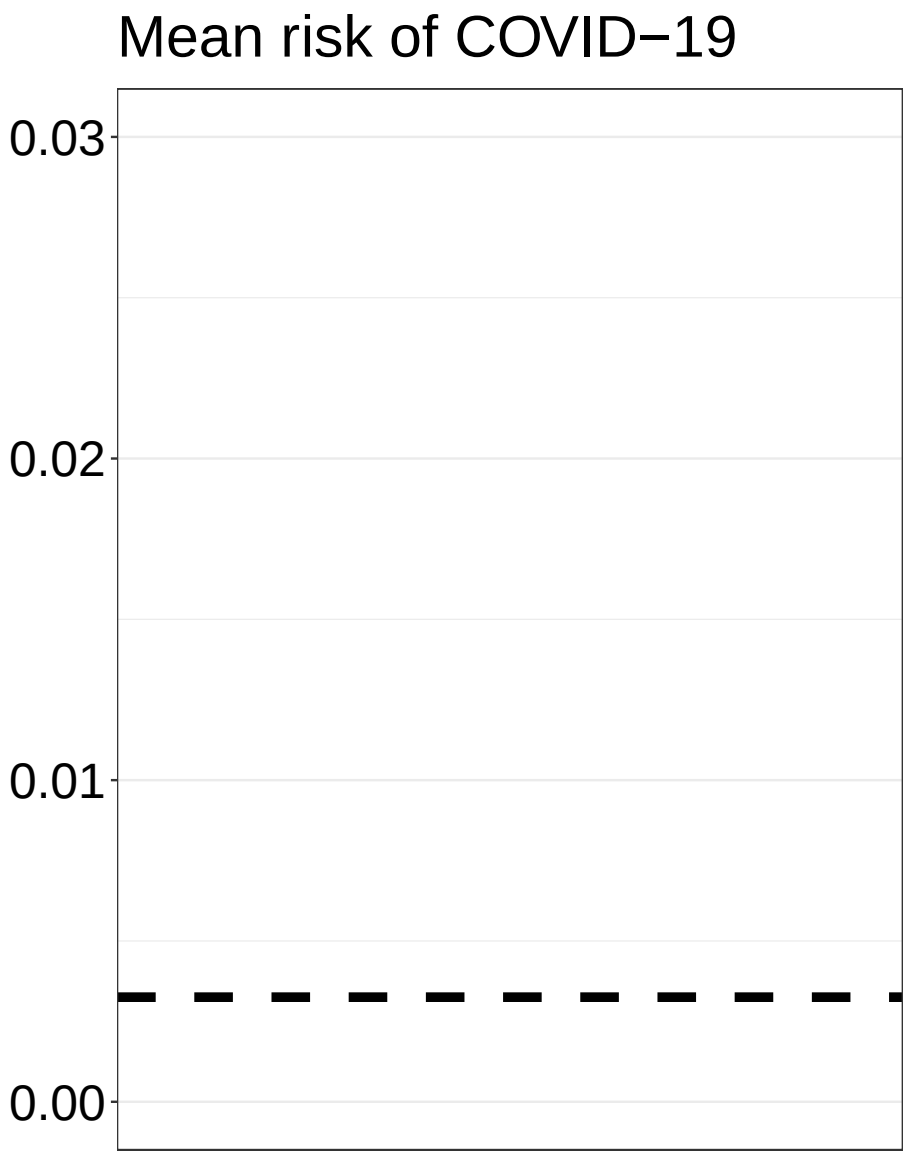
Disease risk under shifted immunogenic responses



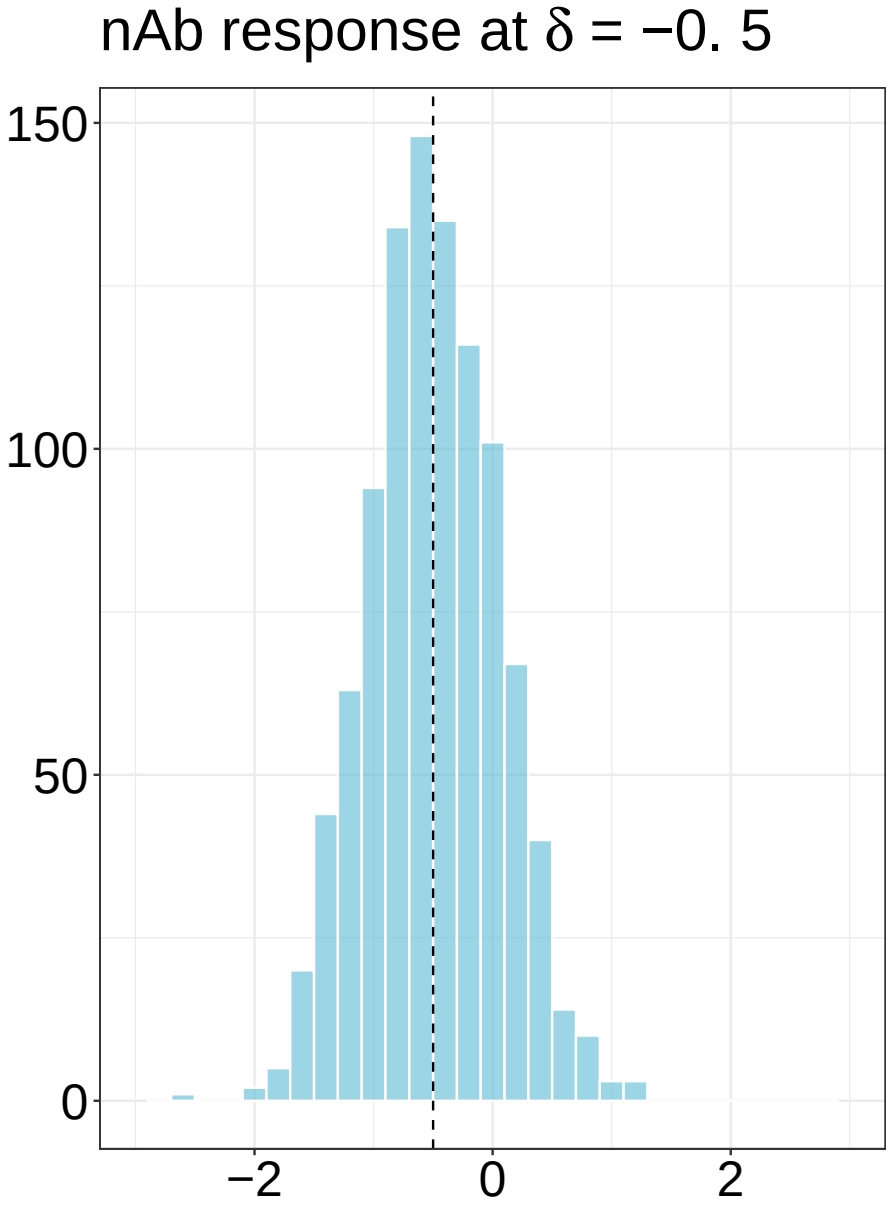


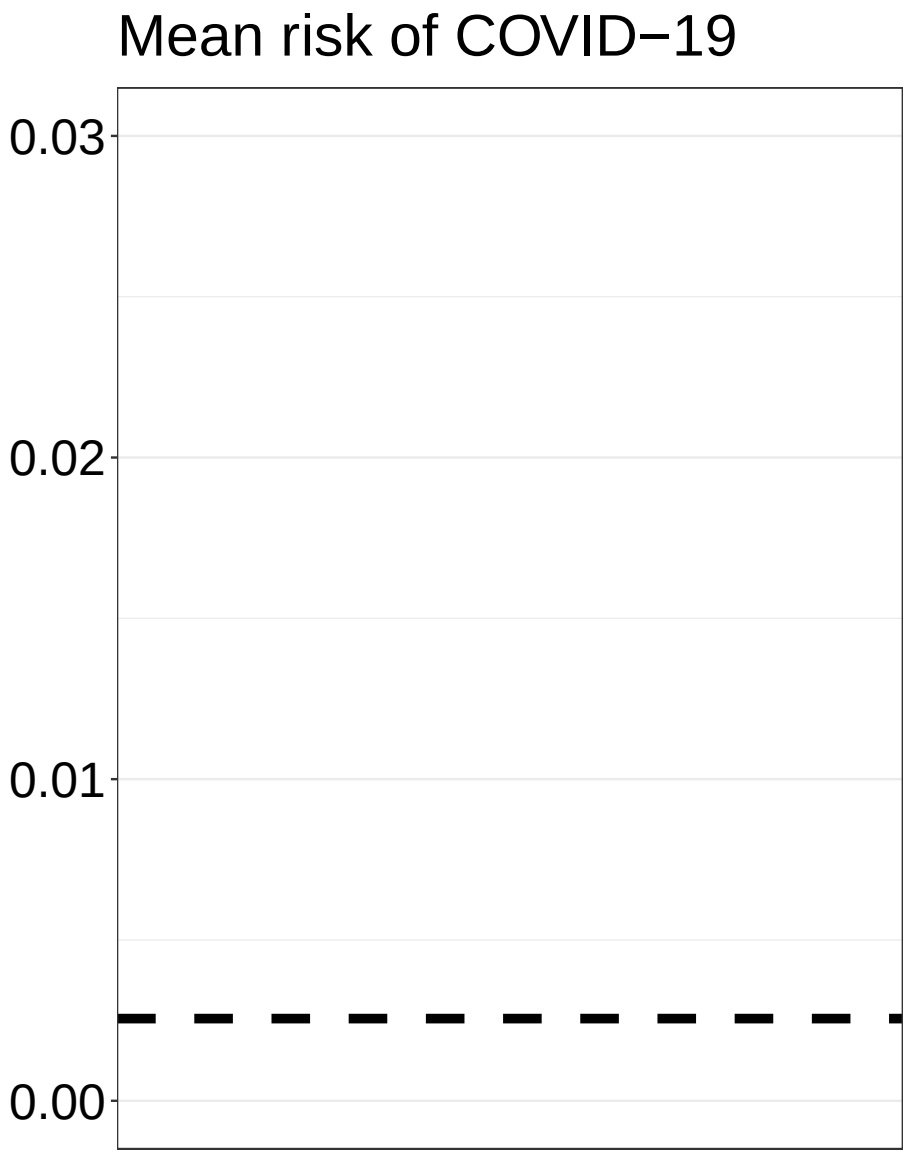
Disease risk under shifted immunogenic responses





Disease risk under shifted immunogenic responses





The causal effects of modified treatment policies

- Stochastic interventions modify the value S would naturally assume by "shifting" the natural exposure distribution.
- Díaz and van der Laan (2012; 2018)'s "shift" interventions[1]

$$d(s, l) = \begin{cases} s + \delta, & s + \delta < u(l) \quad \text{(if plausible)} \\ s, & s + \delta \geq u(l) \quad \text{(otherwise)} \end{cases}$$

- Our estimand is $\psi_{0,\delta} := \mathbb{E}_{\mathrm{P}_{\delta,0}}[Y(d(S, L)) \mid A = 1]$, identified by

$$\psi_{0,\delta} = \int_{\mathcal{L}} \int_{\mathcal{S}} \mathbb{E}_{\mathrm{P}_0}\{Y \mid S = d(s, l), A = 1, L = l\} \\ g_{0,S}(s \mid A = 1, L = l) q_{0,L}(l) d\mu(s) d\nu(l)$$

[1] Haneuse and Rotnitzky (2013) introduced modified treatment policies.

Causal interpretation of statistical target parameter

Assumption 1: *Stable unit treatment value (SUTVA)*

- $Y(d(s_i, l_i))_i$ *does not depend on* $d(s_j, l_j)$ *for* $i = 1, \ldots, n$ *and* $j \neq i$*, or non-interference (Cox 1958, Rubin 1974)*
- $Y(d(s, l)) = Y$ *in the event* $A = 1$, $S = d(s, l)$, *for* $i = 1, \ldots, n$

Assumption 2: *No unmeasured confounding*

$Y(d(s, l)) \perp\!\!\!\perp S \mid A = 1, L = l$, *for* $i = 1, \ldots, n$

Assumption 3: *Structural positivity*

$s \in \mathcal{S} \implies d(s, l) \in \mathcal{S}$ *for all* $l \in \mathcal{L}$, *where* $\mathcal{S}$ *denotes the support of* S *conditional on* $A = 1, L = l$ *for all* $i = 1, \ldots, n$

Interpreting the causal effects of shift interventions

- Consider a data structure: $(Y_s, s \in \mathcal{S})$.
- Let $Y_s = \beta_0 + \beta_1 s + \epsilon_s$, with error $\epsilon_s \sim \mathsf{N}(0, \sigma_s^2)\ \forall\ s \in S$.
- For the counterfactual outcomes $(Y_{s'+\delta}, Y_{s'})$, their difference $Y_{s'+\delta} - Y_{s'}$ may be expressed (for some $s' \in \mathcal{S}$)

$$\begin{aligned}
\mathbb{E}Y_{s'+\delta} - \mathbb{E}Y_{s'} &= [\beta_0 + \beta_1(s' + \delta) + \cancel{\mathbb{E}\epsilon_{s'+\delta}}] - [\beta_0 + \beta_1 s' + \cancel{\mathbb{E}\epsilon_{s'}}] \\
&= \cancel{\beta_0 - \beta_0} + \cancel{\beta_1 s' - \beta_1 s'} + \beta_1 \delta \\
&= \beta_1 \delta
\end{aligned}$$

- A *unit shift* for $s' \in S$ (i.e., for $\delta = 1$) *causes* a counterfactual difference in Y of magnitude β_1 in this simple schematic.

Stochastic-interventional vaccine efficacy (SVE)

- Statistical parameter for vaccine efficacy (VE) estimands:

$$\begin{aligned}\mathrm{SVE}(\delta) &= 1 - \frac{\mathbb{E}[\overline{Q}_{0,Y}(d(S,L),L) \mid A=1]}{\mathbb{P}(Y=1 \mid A=0)} \\ &= 1 - \frac{\psi_{0,\delta}}{\mathbb{P}(Y=1 \mid A=0)} \ ,\end{aligned}$$

  where $\overline{Q}_{0,Y}(s,l) := \mathbb{P}(Y=1 \mid S=s, A=1, L=l)$ is the risk among vaccine recipients.

- $\mathbb{P}(Y=1 \mid A=0)$: counterfactual risk in the placebo arm — i.e., under randomization, $\mathbb{P}(Y=1 \mid A=0) \equiv \mathbb{P}(Y(0)=1)$.

- Summarizes VE via stochastic interventions across δ, per the CoVPN immune correlates SAP[2] (Gilbert et al. 2021a;b).

[2]SAP published at `https://doi.org/10.6084/m9.figshare.13198595`.

Identification: what randomization does (and doesn't) buy

- Causal identification assumptions differ slightly for CVE and SVE, but both require
  1. SUTVA (consistency, no interference);
  2. no unmeasured confounding of the
     - A–S relationship: fulfilled by randomization of A ✓
     - A–Y relationship: fulfilled by randomization of A ✓
     - S–Y relationship: *not* ensured by randomization; needed, conditional on L
- In structural positivity, SVE is more flexible than CVE:
  - CVE: analogous to that of static interventions on A, requiring $\mathbb{P}(S = s \mid A, L) > \epsilon$ for any given s;
  - SVE: based on that of MTPs, that is, $s \in \mathcal{S} \implies d(s, l) \in \mathcal{S}$ for all $l \in \mathcal{L}$.

Estimation in two-phase designs

Estimation of the counterfactual mean $\psi_{0,\delta}$

A *RAL estimator* $\psi_{n,\delta}$ of $\psi_{0,\delta} := \Psi(\mathrm{P}_0)$ is *efficient* if and only if

$$\sqrt{n}(\psi_{n,\delta} - \psi_{0,\delta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} D^{\star}(\mathrm{P}_0)(O_i) + o_P(1) \ ,$$

where $D^{\star}(\mathrm{P})$ is the *efficient influence function* (EIF) of $\psi_{0,\delta}$ with respect to the nonparametric model $\mathcal{P}$ at a distribution $\mathrm{P} \in \mathcal{P}$.

The EIF of $\psi_{0,\delta}$ is indexed by two key *nuisance parameters*

$$\overline{Q}_Y(S, L) := \mathbb{E}_{\mathrm{P}}(Y \mid S, A = 1, L) \qquad \textit{outcome mechanism}$$

$$g_S(S \mid L) := f_{\mathrm{P}}(S \mid A = 1, L) \qquad \textit{generalized propensity score}$$

Flexible, efficient, doubly robust estimation

- Among vaccine recipients ($A = 1$), the efficient influence function of $\psi_{0,\delta}$ with respect to $\mathcal{P}$ is

$$D_F^\star(\mathrm{P}_0)(o) = \frac{g_{0,S}(d^{-1}(s, l) \mid l)}{g_{0,S}(s \mid l)}(y - \overline{Q}_{0,Y}(s, l)) + \overline{Q}_{0,Y}(d(s, l), l) - \psi_{0,\delta}.$$

- The one-step bias-corrected estimator, over the n_1 vaccine recipients:

$$\psi_n^+ = \frac{1}{n_1} \sum_{i:A_i=1} \left[\overline{Q}_{n,Y}(d(S_i, L_i), L_i) + D_{F,n}^\star(O_i)\right] \ .$$

- The TML estimator updates initial estimates of $\overline{Q}_{n,Y}$ to $\overline{Q}_{n,Y}^\star$ via a tilting procedure that sets $\mathrm{P}_{n_1} D_F^\star(\hat{\mathrm{P}}_n^\star) \approx 0$:

$$\psi_n^\star = \frac{1}{n_1} \sum_{i:A_i=1} \overline{Q}_{n,Y}^\star(d(S_i, L_i), L_i) \ .$$

- "Doubly robust" — flexible modeling of nuisance parameters.

Augmented estimators for two-phase sampling designs

- Rose and van der Laan (2011) suggested inverse probability of censoring weighted (IPCW) loss functions:

$$\mathcal{L}(\mathrm{P}_0^X)(O) = \frac{R}{g_{0,R}(Y, L)} \mathcal{L}(\mathrm{P}_0^X)(X)$$

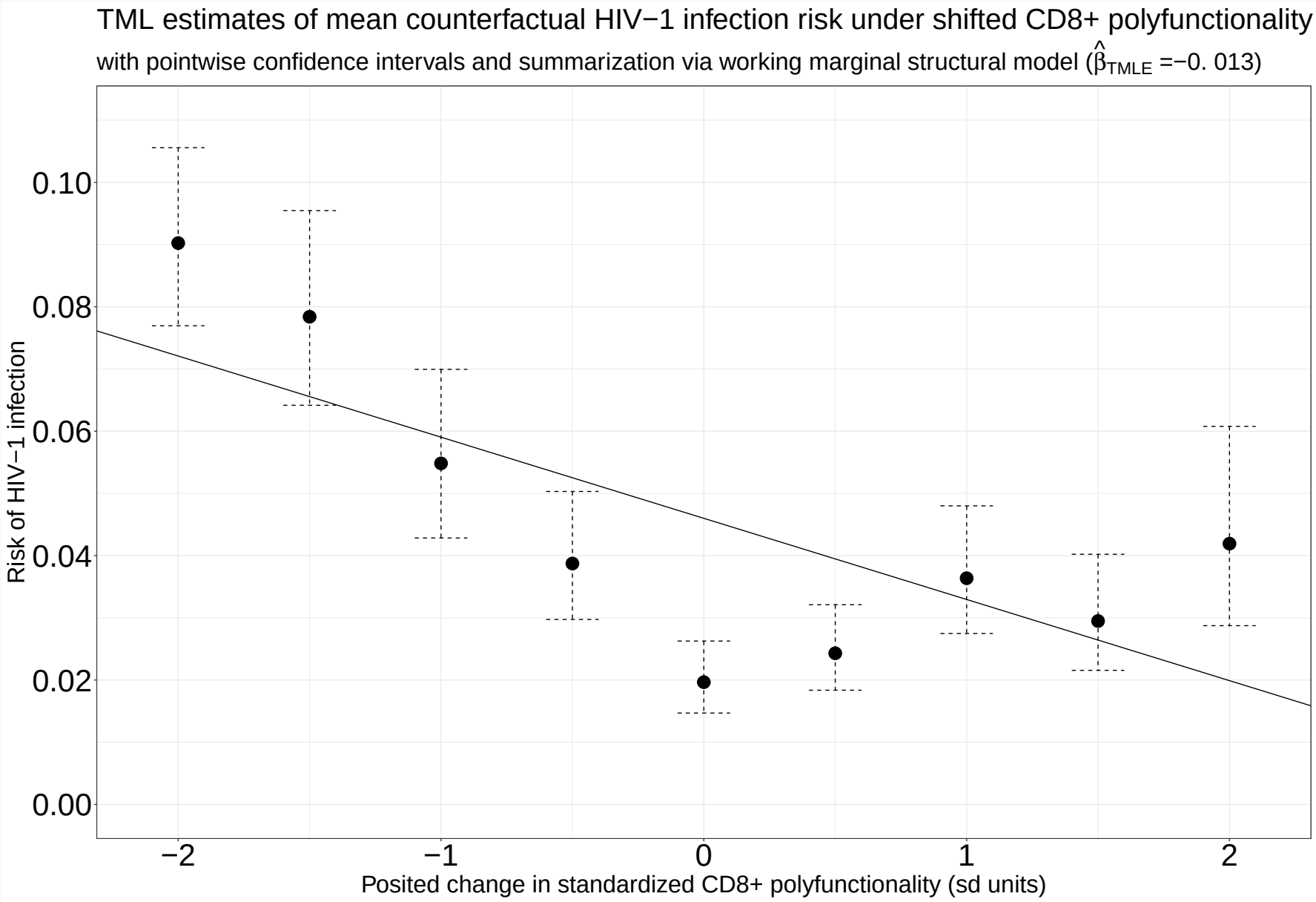
- When the sampling mechanism $g_{0,R}(Y, L)$ is known by design, this procedure yields a reasonably reliable estimator.

- When data-adaptive regression is warranted — that is, when $g_{0,R}(Y, L)$ is not known by design[3]— this is inefficient.

[3]Sampling of non-cases in HVTN 505 used matching (Janes et al. 2017).

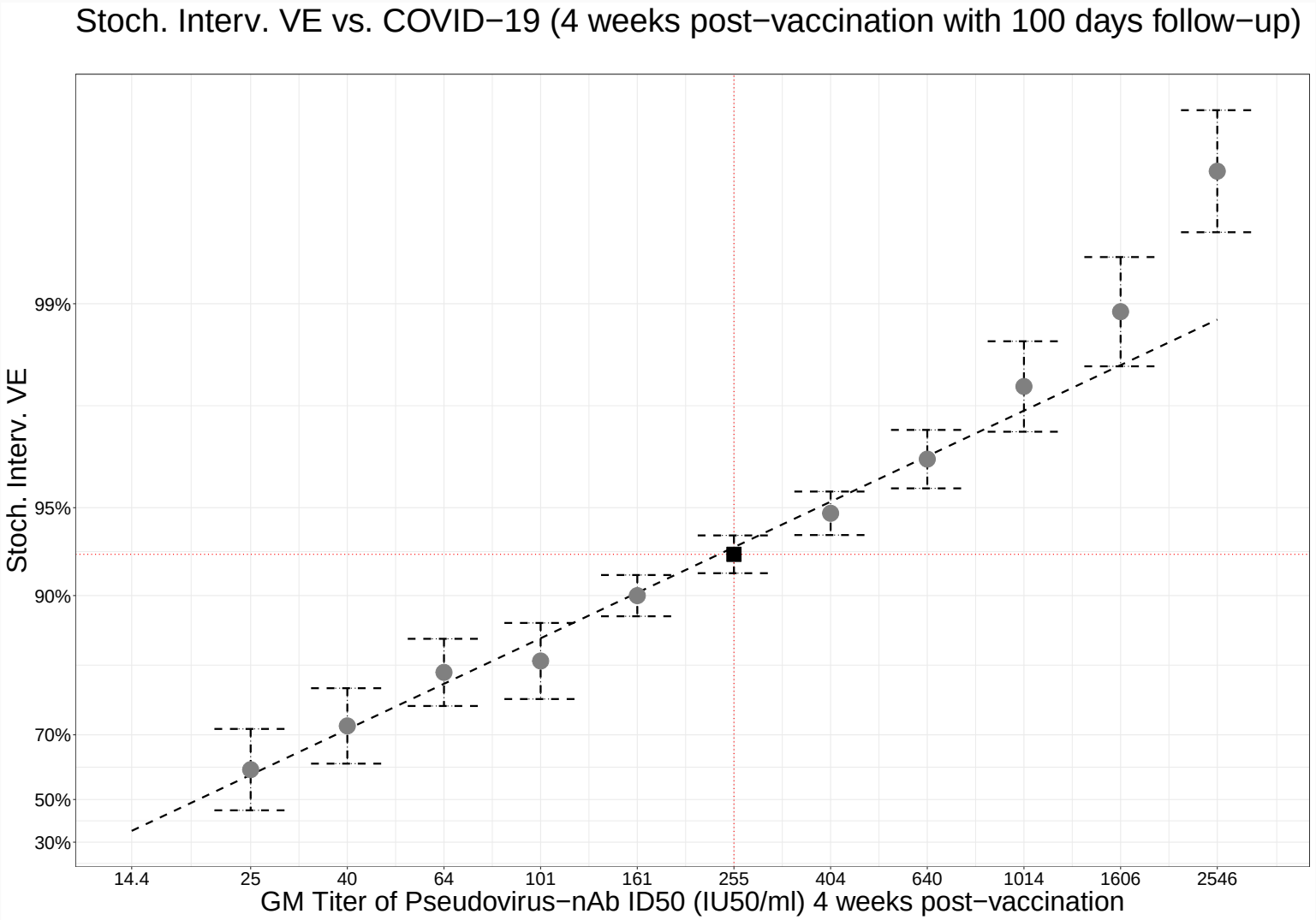
Predicting and bridging vaccine efficacy

SVE *prediction* of HIV-1 risk in the HVTN 505 trial



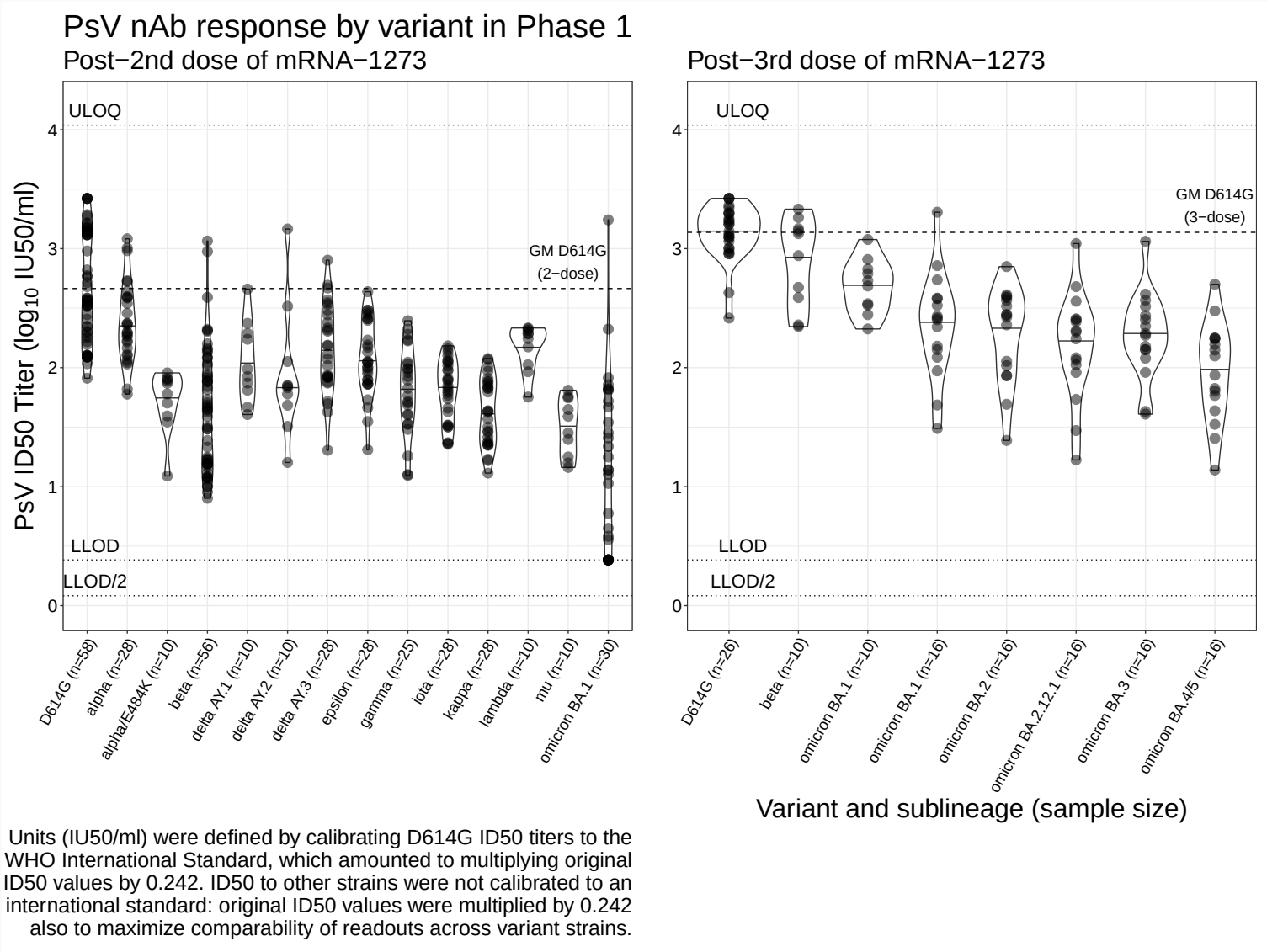
HIV-1 risk change across CD8+ score (`txshift` `R` package).

SVE *prediction* of mRNA-1273 VE in the CoVPN COVE trial



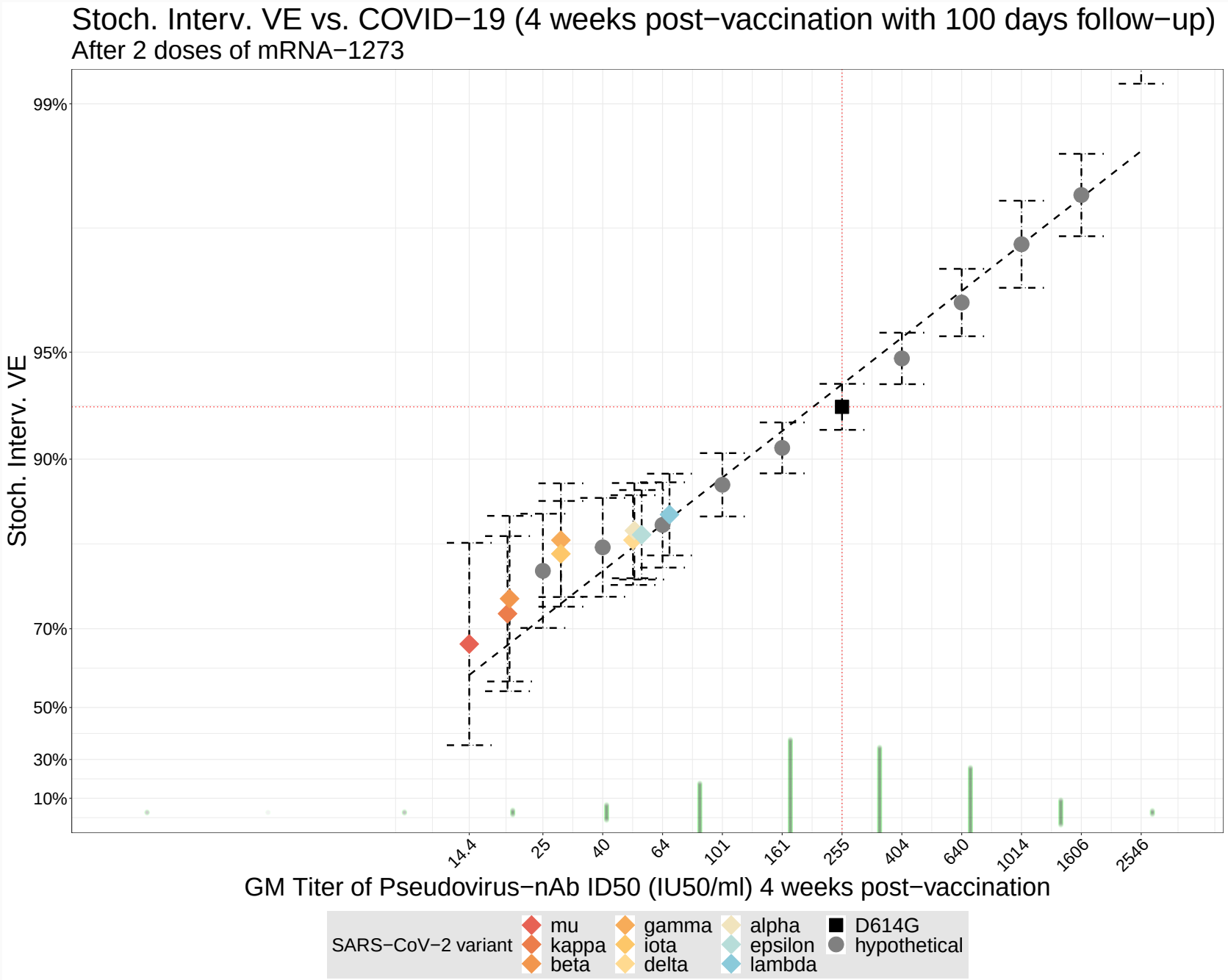
Predicted VE under hypothetical shifts of Day 57 PsV nAb titer (Hejazi et al. 2023, Huang et al. 2023).

Pooled phase 1 studies: PsV nAb responses across variants



Units (IU50/ml) were defined by calibrating D614G ID50 titers to the WHO International Standard, which amounted to multiplying original ID50 values by 0.242. ID50 to other strains were not calibrated to an international standard: original ID50 values were multiplied by 0.242 also to maximize comparability of readouts across variant strains.

SVE *bridging* of mRNA-1273 VE (Hejazi et al. 2023)



The Big Picture

- Flexible stochastic interventions help to formulate novel modified treatment policies (based on "natural" treatment conditions).
- *Modified treatment policies* address causal questions about *realistic* manipulations of quantitative intervention variables.
- Efficient estimators with double/multiple robustness can safely answer such questions *while* incorporating machine learning.
- SVE and CVE are complementary tools for causal mediation analysis of immune correlates (Gilbert et al. 2024).
- Open source software for such statistical analyses is critical for the methods to have any impact on real-world studies.

Causal mediation analysis can provide *causally* grounded evidence for/against an immune marker's value as a surrogate endpoint.

Thank you! Questions?

https://nimahejazi.org/

https://github.com/nshlab

https://github.com/nhejazi

Core methods for SVE (*Biometrics*)	doi:10.1111/biom.13375
SVE application and bridging (*IJID*)	doi:10.1016/j.ijid.2023.09.012
SVE in the COVE trial (*Viruses*)	doi:10.3390/v15102029
Core methods for CVE (*Biostatistics*)	doi:10.1093/biostatistics/kxac024
CoVPN arsenal for CoPs (*Vaccine*)	doi:10.1016/j.vaccine.2024.02.071

Appendix

Immune correlates of protection (Plotkin and Gilbert 2012)

- Correlate of Protection (CoP): immune marker causally predictive of vaccine efficacy, not necessarily mechanistic.
- Mechanistic CoP (mCoP): immune marker that is causally and mechanistically responsible for protection.
- Nonmechanistic CoP (nCoP): immune marker that is predictive but not a mechanistic agent of protection.
- A CoP is a *candidate surrogate* endpoint (Prentice 1989) — primary endpoint in future trials if reliably predictive.

Literature: Díaz and van der Laan (2012; 2018)

- *Proposal:* Evaluate outcome under an altered *intervention distribution* — e.g., $\mathrm{P}_{\delta}(g_{0,S})(S = s \mid L) = g_{0,S}(s - \delta(L) \mid L)$.
- Identification conditions for a statistical parameter of the counterfactual outcome $\psi_{0,\delta}$ under such an intervention.
- Show that the causal quantity of interest $\mathbb{E}_{\mathrm{P}_0^{\delta}}\{Y_{d(S,L)}\}$ is identified by a functional of the distribution of O, i.e.,

$$\psi_{0,\delta} = \int_{\mathcal{L}} \int_{\mathcal{S}} \mathbb{E}_{\mathrm{P}_0}\{Y \mid S = d(s, l), L = l\} \\ g_{0,S}(s \mid L = l) \cdot q_{0,L}(l) d\mu(s) d\nu(l)$$

Literature: Haneuse and Rotnitzky (2013)

- *Proposal:* Characterization of stochastic interventions as *modified treatment policies* (MTPs).
- Assumption of *piecewise smooth invertibility* allows for the post-intervention distribution of any MTP to be recovered:

$$g_{0,S}(s \mid l; \delta) = \sum_{j=1}^{J(l)} \mathbb{I}_{\delta,j}\{h_j(s,l), l\} g_{0,S}\{h_j(s,l) \mid l\} h_j'(s,l)$$

- MTPs account for the natural value of exposure S yet may be interpreted as imposing an altered intervention mechanism.

Slope in a semiparametric model

- Consider the stochastic intervention $g_\delta(\cdot \mid L)$:

$$\begin{aligned}\mathbb{E}Y_{g_\delta} &= \int_L \int_s \mathbb{E}(Y \mid S = s, L) g(s - \delta \mid L) ds dP_0(L) \\ &= \int_L \int_z \mathbb{E}(Y \mid S = z + \delta, L) g(z \mid L) dz dP_0(L),\end{aligned}$$

  defining the change of variable $z = s - \delta$.

- For a semiparametric model, $\mathbb{E}(Y \mid S = z, L) = \beta z + \theta(L)$:

$$\begin{aligned}\mathbb{E}Y_{g_\delta} - \mathbb{E}Y &= \int_L \int_z [\mathbb{E}(Y \mid S = z + \delta, L) - \mathbb{E}(Y \mid S = z, L)] \\ &\qquad g(z \mid L) dz dP_0(L) \\ &= [\beta(z + \delta) + \theta(L)] - [\beta z + \theta(L)] \\ &= \beta\delta\end{aligned}$$

Flexible conditional density estimation of $g_{0,S}$

- Díaz and van der Laan (2011)'s conditional density estimator:

$$g_{n,\alpha}(s \mid L) = \frac{\mathbb{P}(s \in [\alpha_{t-1}, \alpha_t) \mid L)}{\alpha_t - \alpha_{t-1}}.$$

  - Re-expressed as hazard regressions in repeated measures data.
  - Tuning parameter $t \approx$ bandwidth in kernel density estimation.

- When càdlàg (RCLL) with finite sectional variation, we have

$$\text{logit}\{\mathbb{P}(s \in [\alpha_{t-1}, \alpha_t) \mid L)\} = \beta_0 + \sum_{w \subset \{1,\ldots,d\}} \sum_{i=1}^{n} \beta_{w,i} \phi_{w,i},$$

for appropriate basis functions $\{\phi_{w,i}\}_{i=1}^{n}$ (Gill et al. 1995).

Flexible conditional density estimation of $g_{0,S}$

- Utilizing a particular basis construction for ϕ_w, van der Laan (2017)'s HAL estimator achieves $n^{-1/4}$ convergence rate[4].
- Loss-based cross-validation allows selection of a suitable HAL estimator, which has only the ℓ_1 regularization term λ:

$$\beta_{n,\lambda} = \underset{\beta:|\beta_0|+\sum_{w\subset\{1,\ldots,d\}}\sum_{i=1}^{n}|\beta_{w,i}|<\lambda}{\arg\min} \mathrm{P}_n\mathcal{L}(g_{\beta,\lambda,S}),$$

  where $\mathcal{L}(\cdot)$ is an appropriate loss function, giving $\{\lambda_n, \beta_n\}$.

- We denote by $g_{n,\lambda,S} \coloneqq g_{\beta_{n,\lambda},S}$, the HAL estimate of $g_{0,S}$.
- Our `haldensify` `R` package implements our estimator of $g_{0,S}$.

[6]Similar rates can be achieved via *local* (vs. global) smoothness assumptions on $g_{n,S}$ (see, e.g., Robins et al. 2008, Mukherjee et al. 2017, Liu et al. 2021).

A useful class of functions

Consider space of *càdlàg* functions with *finite variation norm*.

Def. càdlàg = *right-hand continuous* with *left-hand limits*

Variation norm Let $\theta_s(u) = \theta(u_s, 0_{s^c})$ be the *section* of θ that sets the coordinates not in s *equal to zero*.

The *variation norm* of θ can be written:

$$|\theta|_v = |\theta(0)| + \sum_{s \subset \{1,\ldots,d\}} \int |d\theta_s(u_s)|,$$

where $x_s = (x(j) : j \in s)$ and the sum is over all subsets.

Sectional variation norm

We can represent the function θ as

$$\theta(x) \quad = \quad \theta(0) + \sum_{s \subset \{1,\ldots,d\}} \int \mathbb{I}(x_s \geq u_s) d\theta_s(u_s),$$

For discrete measures $d\theta_s$ with *support points* $\{u_{s,j} : j\}$ we get a *linear combination* of indicator *basis functions*:

$$\theta(x) = \theta(0) + \sum_{s \subset \{1,\ldots,d\}} \sum_j \beta_{s,j} \theta_{u_{s,j}}(x),$$

where $\beta_{s,j} = d\theta_s(u_{s,j})$, $\theta_{u_{s,j}}(x) = \mathbb{I}(x_s \geq u_{s,j})$, and

$$|\theta|_v = |\theta(0)| + \sum_{s \subset \{1,\ldots,d\}} \sum_j |\beta_{s,j}|.$$

Convergence rate for HAL estimators

We have, for $\alpha(d) = 1/(d+1)$,

$$|\theta_{n,M} - \theta_{0,M}|_{P_0} = O_P(n^{-(1/4+\alpha(d)/8)}).$$

Thus, if we select $M > |\theta_0|_v$, then

$$|\theta_{n,M} - \theta_0|_{P_0} = O_P(n^{-(1/4+\alpha(d)/8)}) \ .$$

Due to oracle inequality for the cross-validation selector M_n,

$$|\theta_{n,M_n} - \theta_0|_{P_0} = O_P(n^{-(1/4+\alpha(d)/8)}) \ .$$

Improved convergence rate (Bibaut and van der Laan 2019):

$$|\theta_{n,M_n} - \theta_0|_{P_0} = O_P(n^{-1/3} \log(n)^{d/2}) \ .$$

Algorithm for TML estimation

1. Construct initial estimators $g_{n,S}$ of $g_{0,S}(S, L)$ and $Q_{n,Y}$ of $\overline{Q}_{0,Y}(S, L)$, perhaps using data-adaptive regression techniques.
2. For each observation i, compute an estimate $H_n(s_i, l_i)$ of the auxiliary covariate $H_0(s_i, l_i)$.
3. Estimate the parameter ϵ in the logistic regression model

$$\text{logit}\overline{Q}_{\epsilon,n,Y}(s, l) = \text{logit}\overline{Q}_{n,Y}(s, l) + \epsilon H_n(s, l),$$

   or an alternative regression model incorporating weights.
4. Compute TML estimator Ψ_n of the target parameter, defining the update $\overline{Q}^{\star}_{n,Y}$ of the initial estimate $\overline{Q}_{n,Y}$ to $\overline{Q}_{\epsilon_n,n,Y}$:

$$\Psi_n = \Psi(\hat{\mathrm{P}}^{\star}_n) = \frac{1}{n}\sum_{i=1}^{n} \overline{Q}^{\star}_{n,Y}(d(S_i, L_i), L_i).$$

Algorithm for IPCW-TML estimation

1. Using all observed units, estimate the sampling mechanism $g_{0,R}(Y, L)$, perhaps using data-adaptive regression methods.
2. Using only observed units in the phase-two sample $R = 1$, construct initial estimators $g_{n,S}(S, L)$ and $\overline{Q}_{n,Y}(S, L)$, weighting by the sampling mechanism estimate $g_{n,R}(Y, L)$.
3. With the approach described for the full data case, compute $H_n(s_i, l_i)$, and fluctuate submodel via logistic regression.
4. Compute IPCW-TML estimator Ψ_n of the target parameter, by solving the IPCW-augmented EIF estimating equation.
5. Iteratively update estimated sampling weights $g_{n,R}(Y, L)$ and IPCW-augmented EIF, updating TMLE in each iteration.

Efficiency and multiple robustness (Hejazi et al. 2021)

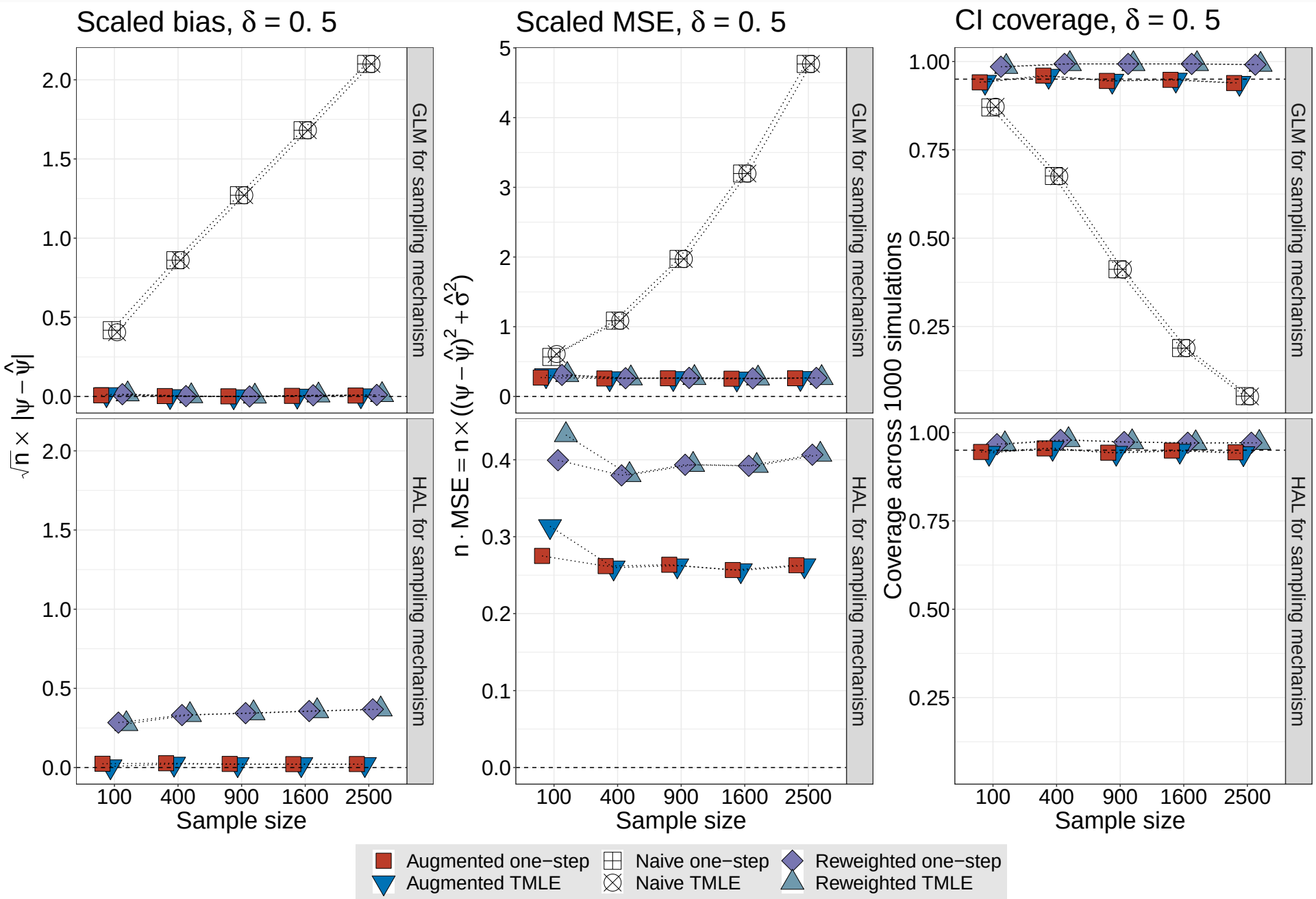
- Then, the IPCW augmentation must be applied to the EIF[5]:

$$D^\star(\mathrm{P}_0^X)(o) = \frac{r}{g_{0,R}(y, l)} D_F^\star(\mathrm{P}_0^X)(x) - \left(1 - \frac{r}{g_{0,R}(y, l)}\right) \mathbb{E}(D_F^\star(\mathrm{P}_0^X)(x) \mid R = 1, Y = y, L = l) \ .$$

- Expresses observed data EIF $D^\star(\mathrm{P}_0^X)(o)$ via complete data EIF $D_F^\star(\mathrm{P}_0^X)(x)$; inclusion of second term improves efficiency.
- An emergent robustness property — one each of $g_{0,S}(S \mid L)$, $\overline{Q}_{0,Y}(S, L)$ and $g_{0,R}(Y, L)$, $\mathbb{E}(D_F^\star(\mathrm{P}_0^X)(x) \mid R = 1, Y, L)$.
- Our `txshift` `R` package implements our estimators of $\psi_{0,\delta}$.

[4] Robins et al. (1994) explored a similar correction for related sampling designs.

Comparing reweighted and augmented estimators



References

Antia, R. and Halloran, M. E. (2021). Transition to endemicity: Understanding COVID-19. *Immunity*, 54(10):2172–2176.

Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S. A., Rouphael, N., Creech, C. B., McGettigan, J., Khetan, S., Segall, N., Solis, J., Brosz, A., Fierro, C., Schwartz, H., Neuzil, K., Corey, L., Gilbert, P., Janes, H., Follmann, D., Marovich, M., Mascola, J., Polakowski, L., Ledgerwood, J., Graham, B. S., Bennett, H., Pajon, R., Knightly, C., Leav, B., Deng, W., Zhou, H., Han, S., Ivarsson, M., Miller, J., Zaks, T., and the COVE Study Group (2021). Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *New England Journal of Medicine*, 384(5):403–416.

Bibaut, A. F. and van der Laan, M. J. (2019). Fast rates for empirical risk minimization over càdlàg functions with bounded sectional variation norm. *arXiv preprint arXiv:1907.09244*.

Breslow, N., McNeney, B., and Wellner, J. A. (2003). Large sample theory for semiparametric regression models with two-phase, outcome dependent sampling. *The Annals of Statistics*, 31(4):1110–1139.

Cox, D. R. (1958). *Planning of Experiments*. Wiley.

Díaz, I. and van der Laan, M. J. (2011). Super learner based conditional density estimation with application to marginal structural models. *International Journal of Biostatistics*, 7(1):1–20.

Díaz, I. and van der Laan, M. J. (2012). Population intervention causal effects based on stochastic interventions. *Biometrics*, 68(2):541–549.

Díaz, I. and van der Laan, M. J. (2018). Stochastic treatment regimes. In *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*, pages 167–180. Springer Science & Business Media.

Follmann, D. (2006). Augmented designs to assess immune response in vaccine trials. *Biometrics*, 62(4):1161–1169.

Gilbert, P. B., Fong, Y., Benkeser, D., Andriesen, J., Borate, B., Carone, M., Carpp, L. N., Díaz, I., Fay, M. P., Fiore-Gartland, A., Hejazi, N. S., Huang, Y., Huang, Y., Hyrien, O., Janes, H. E., Juraska, M., Li, K., Luedtke, A., Nason, M., Randhawa, A. K., van der Laan, L., Williamson, B. D., Zhang, W., and Follmann, D. (2021a). CoVPN COVID-19 vaccine efficacy trial immune correlates statistical analysis plan.

Gilbert, P. B., Fong, Y., Hejazi, N. S., Kenny, A., Huang, Y., Carone, M., Benkeser, D., and Follmann, D. (2024). Four statistical frameworks for assessing an immune correlate of protection (surrogate endpoint) from a randomized, controlled, vaccine efficacy trial. *Vaccine*, 42(9):2181–2190.

Gilbert, P. B., Fong, Y., Kenny, A., and Carone, M. (2023). A controlled effects approach to assessing immune correlates of protection. *Biostatistics*, 24(4):850–865.

Gilbert, P. B., Montefiori, D. C., McDermott, A. B., Fong, Y., Benkeser, D., Deng, W., Zhou, H., Houchens, C. R., Martins, K., Jayashankar, L., Castellino, F., Flach, B., Lin, B. C., O'Connell, S., McDanal, C., Eaton, A., Sarzotti-Kelsoe, M., Lu, Y., Yu, C., Borate, B., van der Laan, L. W., Hejazi, N. S., Huynh, C., Miller, J., El Sahly, H. M., Baden, L. R., Baron, M., De La Cruz, L., Gay, C., Kalams, S., Kelley, C. F., Kutner, M., Andrasik, M. P., Kublin, J. G., Corey, L., Neuzil, K. M., Carpp, L. N., Pajon, R., Follmann, D., Donis, R. O., Koup, R. A., and on behalf of the Immune Assays; Moderna, Inc.; Coronavirus Vaccine Prevention Network (CoVPN) / Coronavirus Efficacy (COVE); and United States Government (USG) / CoVPN Biostatistics Teams (2021b). Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science*, 375(6576):43–50.

Gill, R. D., van der Laan, M. J., and Wellner, J. A. (1995). Inefficient estimators of the bivariate survival function for three models. In *Annales de l'IHP Probabilités et statistiques*, volume 31, pages 545–597.

Hammer, S. M., Sobieszczyk, M. E., Janes, H., Karuna, S. T., Mulligan, M. J., Grove, D., Koblin, B. A., Buchbinder, S. P., Keefer, M. C., Tomaras, G. D., et al. (2013). Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. *New England Journal of Medicine*, 369(22):2083–2092.

Haneuse, S. and Rotnitzky, A. (2013). Estimation of the effect of interventions that modify the received treatment. *Statistics in Medicine*, 32(30):5260–5277.

Hejazi, N. S., Shen, X., Carpp, L. N., Benkeser, D., Follmann, D., Janes, H. E., Baden, L. R., El Sahly, H. M., Deng, W., Zhou, H., Leav, B., Montefiori, D. C., and Gilbert, P. B. (2023). Stochastic interventional approach to assessing immune correlates of protection: Application to the COVE mRNA-1273 vaccine trial. *International Journal of Infectious Diseases*, 137:28–39.

Hejazi, N. S., van der Laan, M. J., Janes, H. E., Gilbert, P. B., and Benkeser, D. C. (2021). Efficient nonparametric inference on the effects of stochastic interventions under two-phase sampling, with applications to vaccine efficacy trials. *Biometrics*.

Huang, Y., Hejazi, N. S., Blette, B., Carpp, L. N., Benkeser, D., Montefiori, D. C., McDermott, A. B., Fong, Y., Janes, H. E., Deng, W., Zhou, H., Houchens, C. R., Martins, K. A., Jayashankar, L., Flach, B., Lin, B. C., O'Connell, S., McDanal, C., Eaton, A., Sarzotti-Kelsoe, M., Lu, Y., Yu, C., Kenny, A., Carone, M., Huynh, C., Miller, J., El Sahly, H. M., Baden, L. R., Jackson, L. A., Campbell, T. B., Clark, J. L., Andrasik, M. P., Kublin, J. G., Corey, L., Neuzil, K. M., Pajon, R., Follmann, D. A., Donis, R. O., Koup, R. A., Gilbert, P. B., and on behalf of the Immune Assays Team; Moderna, Inc., Team; Coronavirus Prevention Network (CoVPN)/Coronavirus Efficacy (COVE) Team; and the United States Government (USG)/CoVPN Biostatistics Team (2023). Stochastic interventional vaccine efficacy and principal surrogate analyses of antibody markers as correlates of protection against symptomatic COVID-19 in the COVE mRNA-1273 trial. *Viruses*, 15(10):2029.

Janes, H. E., Cohen, K. W., Frahm, N., De Rosa, S. C., Sanchez, B., Hural, J., Magaret, C. A., Karuna, S., Bentley, C., Gottardo, R., et al. (2017). Higher T-cell responses induced by DNA/rAd5 HIV-1 preventive vaccine are associated with lower HIV-1 infection risk in an efficacy trial. *The Journal of Infectious Diseases*, 215(9):1376–1385.

Liu, L., Mukherjee, R., Robins, J. M., and Tchetgen Tchetgen, E. (2021). Adaptive estimation of nonparametric functionals. *Journal of Machine Learning Research*, 22(99):1–66.

Mukherjee, R., Newey, W. K., and Robins, J. M. (2017). Semiparametric efficient empirical higher order influence function estimators. *arXiv preprint arXiv:1705.07577*.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Plotkin, S. A. and Gilbert, P. B. (2012). Nomenclature for immune correlates of protection after vaccination. *Clinical Infectious Diseases*, 54(11):1615–1617.

Prentice, R. L. (1986). A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika*, 73(1):1–11.

Prentice, R. L. (1989). Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in Medicine*, 8(4):431–440.

Robins, J. M., Li, L., Tchetgen Tchetgen, E., and van der Vaart, A. W. (2008). Higher order influence functions and minimax estimation of nonlinear functionals. In *Probability and statistics: essays in honor of David A. Freedman*, pages 335–421. Institute of Mathematical Statistics.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, 89(427):846–866.

Rose, S. and van der Laan, M. J. (2011). A targeted maximum likelihood estimator for two-stage designs. *The International Journal of Biostatistics*, 7(1):1–21.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701.

van der Laan, M. J. (2017). A generally efficient targeted minimum loss based estimator based on the highly adaptive lasso. *International Journal of Biostatistics*, 13(2).