How should we select test-negative controls?

A causal perspective in the era of multiplex respiratory testing

Christopher B. Boyer

Department of Quantitative Health Sciences
Cleveland Clinic Lerner College of Medicine, Case Western Reserve University

Society for Epidemiologic Research • June 24, 2026

Acknowledgements

Collaborators: Kendrick Quijin Li, Xu Shi, Eric Tchetgen Tchetgen, Marc Lipsitch

Conflicts of Interest: I have no conflicts of interest.

Funding: This work was supported by NIH grants R01GM139926 and U01CA261277.

Disclaimer: The content presented here represents the views of the author and not necessarily those of the NIH.

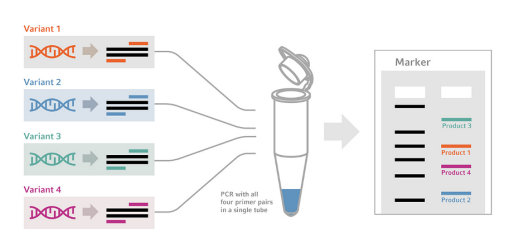
The new diagnostic landscape

A single **multiplex PCR** swab now resolves a dozen respiratory pathogens at once.

For the test-negative design, this has important implications:

- "Test-negative" is no longer one thing...
- ...it is a **mixture** of specific pathogens.

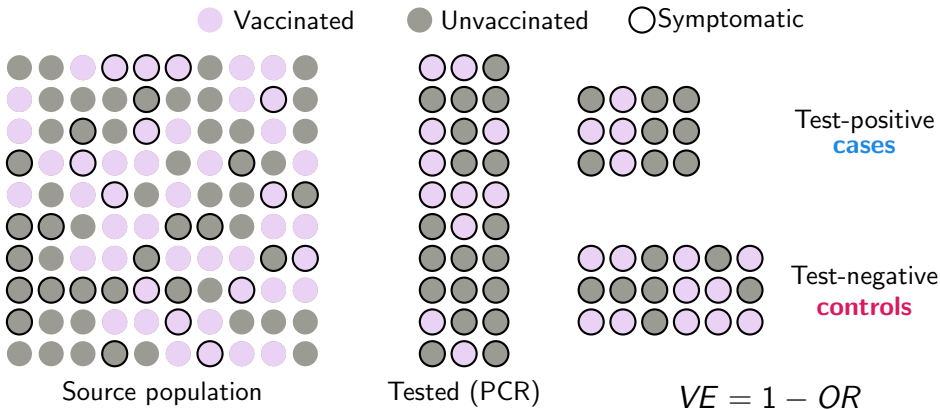
Which of them should be controls?



Source: S. Rosel 2021

The test-negative design

Among the symptomatic who get tested, compare **vaccination odds** in cases vs. controls.



The control group was always a black box

Historically: "influenza-negative influenza-like illness." One *composite* outcome.

The question we previously ignored

If a control is symptomatic but *not* the focal pathogen — what *are* they sick with, and does it matter for causal validity?

Multiplex PCR finally lets us look inside the box. So:

- Are all test-negative pathogens **valid** controls? (No.)
- If not, which do we keep and how do we combine them?
- Can the extra information *improve* the design rather than just complicate it?

Two views on what a control is for

The identification literature contains two answers.

View 1 — Sampling

Controls represent the **source population's** vaccination distribution.

Basis: no unmeasured confounding + control exchangeability (Schnitzer et al. 2022).

Multiplex used *defensively*: **exclude** clearly invalid pathogens.

View 2 — Bias correction

Controls are a **negative-control outcome** (NCO) that can be used for bias correction.

Basis: equi-confounding + valid NCO (Boyer et al. 2026).

Multiplex used *constructively*: **choose** pathogens that mimic the bias.

Shared requirement: **vaccination must not affect the control outcome.**

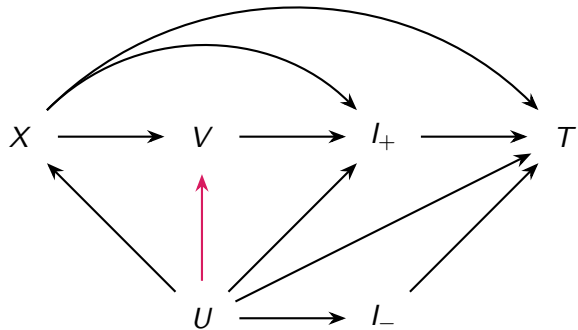
A causal model for the TND

- V: focal vaccination
- I_+: test-positive illness
- I_-: test-negative illness
- T: clinically tested
- X, U: measured / unmeasured

Outcome Y (medically attended):

$$Y = \begin{cases} +1 & I_+ = 1,\ T = 1 \\ -1 & I_- = 1,\ T = 1 \\ 0 & T = 0 \end{cases}$$

The TND conditions on $T = 1$ — a **collider**.



The **red** $U \to V$ arrow highlights unmeasured confounding pathway.

Equi-confounding: controls as an NCO

Decomposition:

$$RR_V(X) = \frac{\Pr(Y^{v=1}=1 \mid V{=}1, X)}{\Pr(Y^{v=0}=1 \mid V{=}1, X)} = \underbrace{\frac{\Pr[Y=1|V=1,X]}{\Pr[Y=1|V=0,X]}}_{\text{observed risk ratio}} \times \underbrace{\frac{\Pr[Y^{v=0}=1|V=0,X]}{\Pr[Y^{v=0}=1|V=1,X]}}_{\text{de-biasing term}}.$$

Assumptions

(A1) Consistency

(A2) No vaccine effect on I_-

(A3) **Equi-confounding** (mult. scale)

(A4) Overlap

A2 implies no effect of V on I_-.

A3 implies effect of U on I_+, I_-, and T is the *same on the multiplicative scale*.

Result — A2 + A3 implies RR for $Y = -1$ equals the de-biasing term. Under TND sampling, odds ratio identifies causal RR.

Write $p_{vy} = \Pr(Y{=}y \mid T{=}1, V{=}v, X)$. Then

$$RR_V(X) = \frac{p_{1,+1}\, p_{0,-1}}{p_{1,-1}\, p_{0,+1}} = OR_{T=1}(X)$$

Same logistic regression. Weaker assumption.

Source: Boyer et al. 2026

Multiplex splits the control node

Test-negative illness is a *union*:

$$I_{-} = \mathbb{1}\Big\{ \bigcup_{k=1}^{K} I_{-k} = 1 \Big\}$$

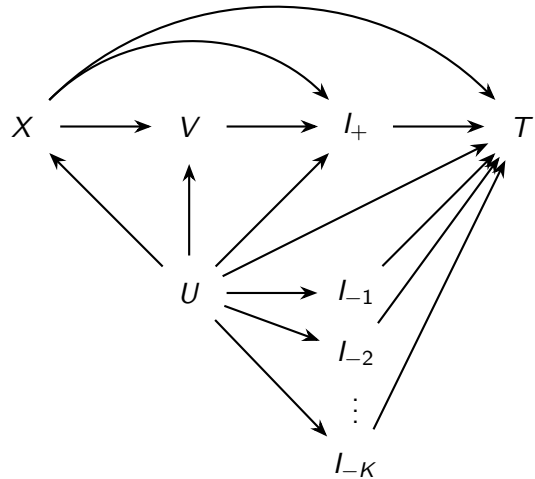
The pooled odds ratio is a prevalence-weighted average of pathogen-specific contrasts:

$$OR_{T=1} \;=\; \sum_{k=1}^{K} w_k \, OR_k$$

The catch

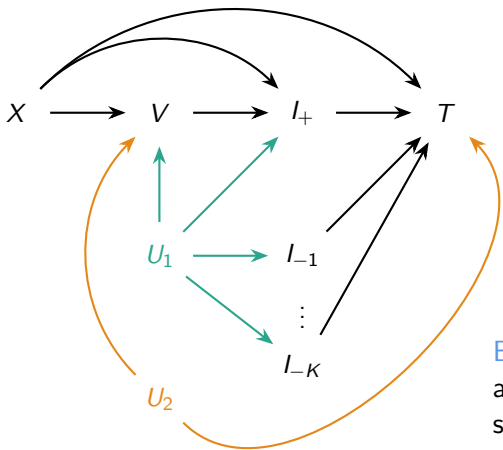
One bad pathogen biases the *whole* pool.

Naive pooling **hides** it.



What makes a good negative control?

A useful pathogen mimics the bias it is meant to cancel. Two flavors of bias:

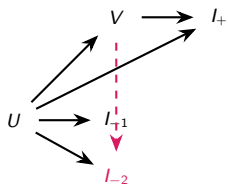


- **Exposure proxies** (U_1): share *how you get infected* — transmission route, crowding, contacts.
  e.g. rhinovirus, parainfluenza for influenza.
- **Testing proxies** (U_2): share *why you get tested* — care-seeking, access.
  e.g. enteric / arboviral febrile illness.

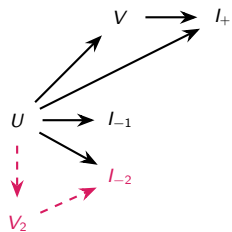
Best controls do both — respiratory viruses with a similar clinical presentation. The taxonomy says *why* they've always "felt right."

Three examples: reasoning about pathogen screening

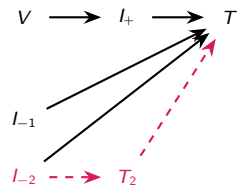
Dashed arrows = the offending pathway (I_{-2} is the problem child).



1. Vaccine irrelevance
Focal vaccine affects a control pathogen. **Exclude it.**



2. Other-vaccine entanglement
Control has its own correlated vaccine. **Exclude / adjust.**

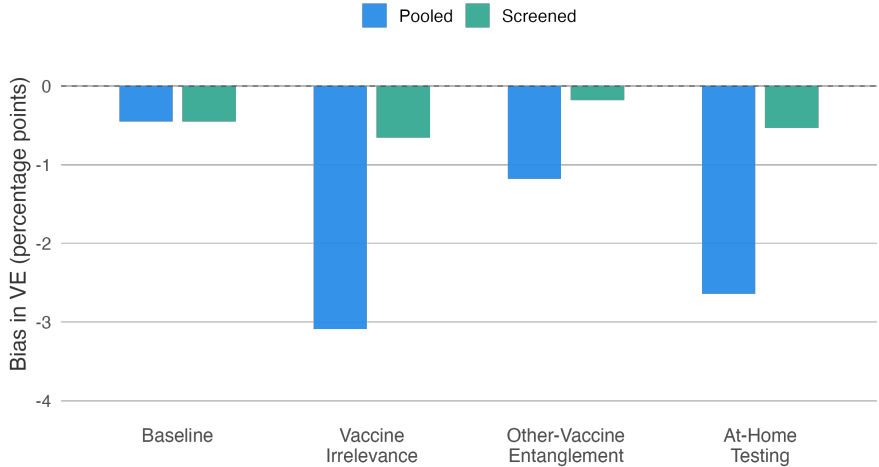


3. Testing comparability
At-home tests reroute care-seeking. **Exclude / adjust.**

All three: a pathway makes one pathogen behave differently from the rest.

Does one bad pathogen really matter?

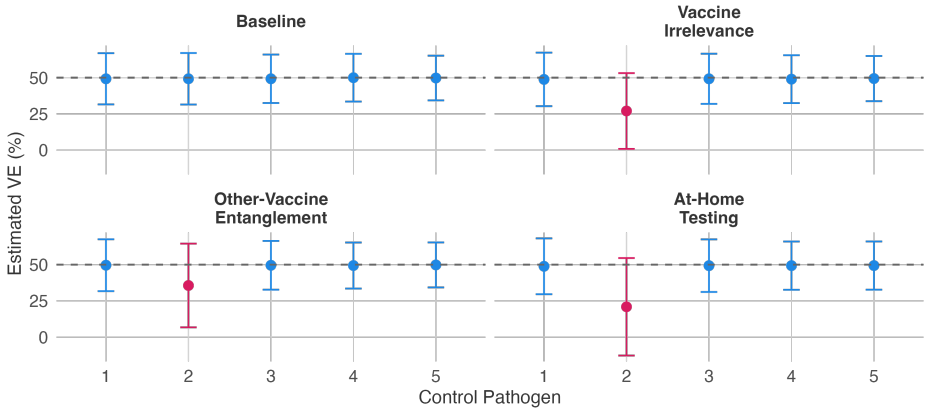
Cohort n=10,000; true VE=50%; K=5 control pathogens; the violation hits *one* of them.



Pooled (all 5) can bias VE by 2–3 pp; **screened** (drop the offender) restores near-nominal performance.

Multiplex also gives you the diagnostic

Pathogen-specific VE estimates: identifying the culprit (**pathogen 2**).



A formal *homogeneity test* across pathogens flags violations — a screening signal you only get with multiplex data.

Also in the paper

The same framework handles the messy realities multiplex panels create:

- **Co-detections.** Control–control: pooled OR is *invariant*. Focal–control: needs a difference-in-differences estimator.
- **Pan-negatives.** When/why to include symptomatic, all-negative episodes.
- **Imperfect assays.** Identification under non-differential sensitivity.
- **Connections.** Reason-for-testing, platform TNDs, proximal inference.
- **Workflow.** A pre-specified control-selection recipe.

Backup slides for any of these.

Takeaways

1. Multiplex testing turns control selection from an afterthought into a causal design choice.
2. View control selection as choosing pontential **negative-control outcomes**: pick pathogens that *mimic the bias*.
3. **Pool carefully** — Pooling reduces variance, but one assumption-violating pathogen can bias the whole estimate (**bias-variance trade off**).
4. Causal framework can help use think critically about screening: e.g., **vaccine irrelevance, other-vaccine/intervention entanglement, testing comparability**.

The control group is not just a sampling convenience. With multiplex data, we can finally treat it like the design choice it always was.

Thank you

Christopher B. Boyer
boyerc5@ccf.org • christopherbboyer.com

Collaborators: Kendrick Quijin Li, Xu Shi, Eric Tchetgen Tchetgen, Marc Lipsitch

Backup — Equi-confounding, factored

(A3) factors into two interpretable pieces:

(A) Infection

U acts on *getting symptomatically infected* the same (multiplicatively) whether the cause is I_+ or I_-:

$$\frac{\Pr(I_+^0{=}1 \mid V{=}1, X)}{\Pr(I_+^0{=}1 \mid V{=}0, X)} = \frac{\Pr(I_-^0{=}1 \mid V{=}1, X)}{\Pr(I_-^0{=}1 \mid V{=}0, X)}$$

(B) Testing

U acts on *seeking a test when symptomatic* the same whether the illness is I_+ or I_-.

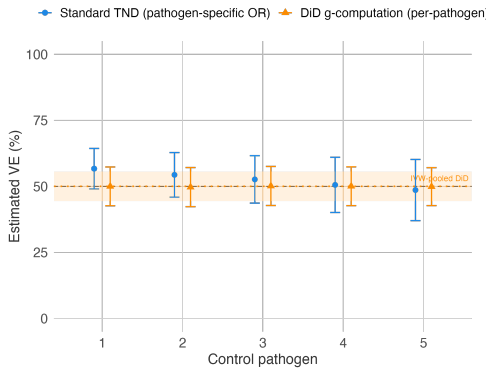
Exposure proxies target (A); testing proxies target (B).

Backup — Focal–control co-detection

When the focal pathogen co-occurs with a control and is assigned to the case group:

- Unvaccinated are more often co-infected ⇒ siphoned into cases.
- Residual control pool becomes *more vaccinated*.
- Pooled TND **overstates** VE — and re-labeling can't fix it.

A **difference-in-differences** (g-computation) estimator on the log scale cancels the common confounding and recovers VE.

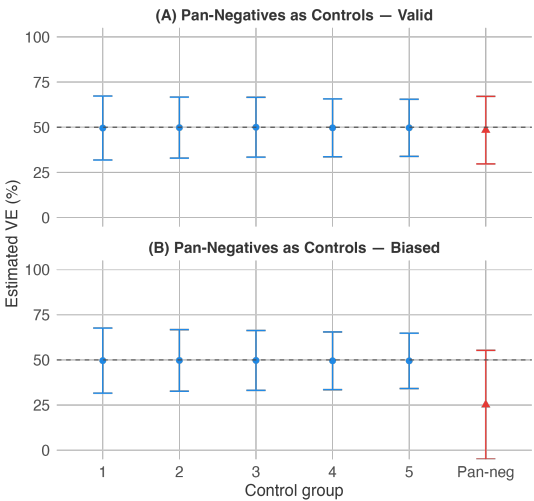


Backup — Pan-negative episodes

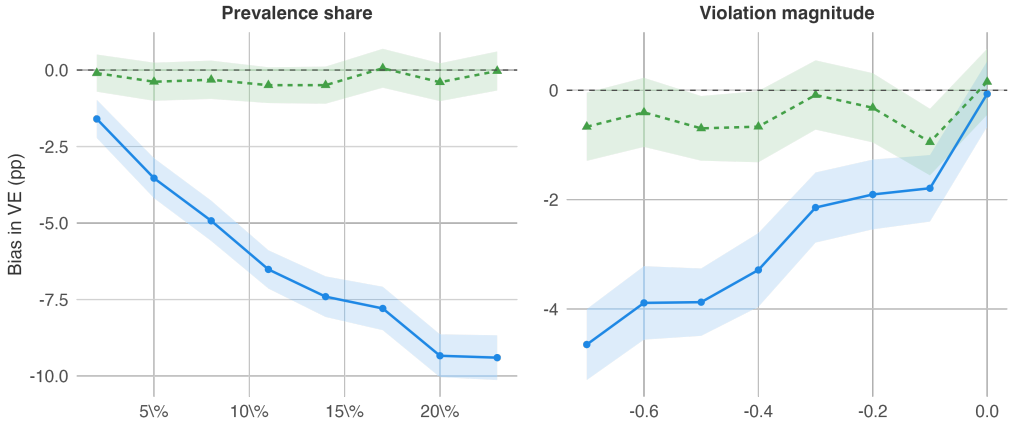
Symptomatic, negative for *everything* on the panel. Could be:

- off-panel pathogens,
- measurement error (false negatives),
- non-infectious symptoms.

Under perfect specificity & non-differential sensitivity, excluding them still identifies RR_V. We recommend reporting *with and without* as a pre-specified sensitivity analysis.



Backup — Bias grows with prevalence & severity



Pooled bias scales with the offending pathogen's share of controls and the magnitude of the violation; screening stays flat near zero.

References I

- Lewnard, Joseph A et al. (2018). "Measurement of Vaccine Direct Effects Under the Test-Negative Design". *American Journal of Epidemiology* 187.12, pp. 2686–2697.
- Li, Kendrick Qijun et al. (2023). "Double Negative Control Inference in Test-Negative Design Studies of Vaccine Effectiveness". *Journal of the American Statistical Association*, pp. 1–12.
- Schnitzer, Mireille E. (2022). "Estimands and Estimation of COVID-19 Vaccine Effectiveness Under the Test-Negative Design". *Epidemiology* 33.3, pp. 325–333.
- Sofer, Tamar et al. (2016). "On Negative Outcome Control of Unobserved Confounding as a Generalization of Difference-in-Differences". *Stat Sci* 31.3, pp. 348–361.
- Sullivan, Sheena G., Eric J. Tchetgen Tchetgen, and Benjamin J. Cowling (2016). "Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness". *American Journal of Epidemiology* 184.5, pp. 345–353.