Scientific Research Publishing

Open Journal of Epidemiology, 2025, 15(3), 628-636
https://www.scirp.org/journal/ojepi
ISSN Online: 2165-7467
ISSN Print: 2165-7459

Studying the Imprinting Effect in Cohort Studies of Covid Vaccines

Eyal Shahar

Department of Epidemiology & Biostatistics, Mel & Enid Zuckerman College of Public Health, University of Arizona, Tucson, USA

Email: shahar@arizona.edu

How to cite this paper: Shahar, E. (2025) Studying the Imprinting Effect in Cohort Studies of Covid Vaccines. *Open Journal of Epidemiology*, **15**, 628-636.
https://doi.org/10.4236/ojepi.2025.153041

Received: July 16, 2025
Accepted: August 12, 2025
Published: August 15, 2025

Abstract

Immune imprinting denotes a phenomenon where previous exposure to a viral antigen, through infection or vaccination, imprints that antigen in immunological memory. When the immune system is exposed to a new strain or a new variant, part of the response targets the historical antigen(s), and the fight against the new offender is less effective. It was recently argued that observational studies that show evidence of immune imprinting by prior vaccination reflect colliding (selection) bias, not imprinting. The collider is a first infection, and the two causes are vaccination and susceptibility to infection. As will be explained, the claim is not uniformly valid, nor can it be verified or refuted based on a construct called susceptibility. Moreover, if the bias is significant at all, underestimation of the true effect of immune imprinting is equally possible.

Keywords

Immune Imprinting, Colliding Bias, Selection Bias, Interaction

1. Introduction

Of the six categories of biases [1], colliding (selection) bias is probably among the least understood. It is the antithetical counterpart of confounding bias, so the pair should be similarly named: confounding bias and colliding bias, confounder bias and collider bias, or confounder non-stratification bias and collider-stratification bias. Selection bias is a historical misnomer that does not cover all the mechanisms by which the bias arises and is sometimes confused with effect-modification bias [2].

Colliding bias might arise following conditioning on a shared effect of two causes. The simplest form of conditioning is restricting the outcome variable to a

single value. When turned into a value, the outcome is dissociated from its causes. However, if the two causes were marginally independent before conditioning on their shared outcome, they will be associated under certain conditions. Similarly, if they were marginally associated, the association would be altered. In either case, an artificial, non-causal component is added to associations that arise from the natural causal structure. This addition might be a source of bias when estimating effects, whether null or otherwise.

Two methodological articles have examined the role of colliding bias in observational studies of immune imprinting [3] [4], a phenomenon that was hypothesized long ago and has been studied for many years [5]. Informally explained, previous exposure to a viral antigen, through infection or vaccination, imprints that antigen in immunological memory. When the immune system is exposed to a new strain or a new variant, part of the response is "mistakenly" targeting the historical antigen(s), and the fight against the new offender is less effective. Imprinting does not wane over time. When first proposed, the phenomenon was called "the original antigenic sin".

The authors argue that observational studies that show evidence of imprinting by prior vaccination reflect colliding bias, not imprinting. The collider is a first infection, and the two causes are vaccination and susceptibility to infection. As will be explained, the claim is not uniformly valid. Moreover, it cannot be verified or refuted based on a construct called "susceptibility", which does not really belong in a causal diagram. A person has a set of attributes—natural variables—that might affect getting infected, but susceptibility is not an attribute. Rather, it is the output of an unspecified function of unspecified variables, similar to constructs like assertiveness and attractiveness. It will be displayed, however, and shown to be insufficient.

2. The Causal Structure

Figure 1 replicates the causal structure that was displayed in one of the articles

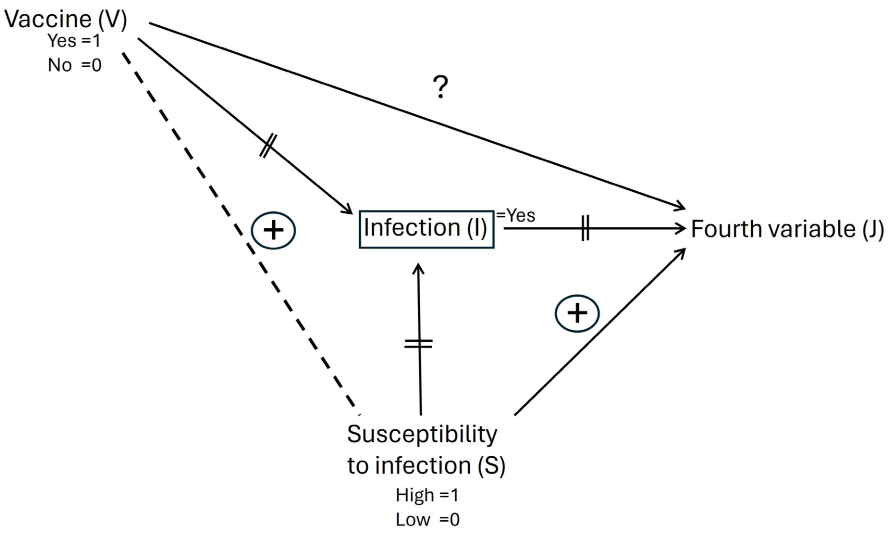

Figure 1. Figure 1c in the article in a different layout and notation.

[4], in a different layout and notation. V is vaccination status, S is susceptibility status, and I is infection status. The fourth variable, J, is a second infection or some other outcome of infection. To simplify, consider a binary version of susceptibility to infection: high or low.

A frame indicates conditioning. Two lines over an arrow denote dissociation between the two ends, despite a causal arrow, because one end is transformed from a variable (I) to a value (I = *yes*). The dashed line explicitly depicts an induced association. It was displayed in the first landmark publication on causal diagrams in the epidemiological literature [6]. Unfortunately, it is usually omitted. The arrow with the question mark is the effect of interest: Is it positive, negative, or null?

In either case, the association between V and J does not provide the correct answer—qualitatively or quantitatively—if the dashed line exists. The path V---S → J adds a non-causal component—colliding bias. That's a valid conclusion of the articles.

3. Approach

The direction of the induced association, and hence the direction of the bias, depends on the direction of the interaction between V and S [7] [8], which can be converted to the direction of effect modification. The last sentence applies to the magnitude of the association as well [7] [8]. Therefore, even if the bias exists in the stated direction, the problem might be overestimation of a true effect. It is not necessarily an all-or-none situation.

Three questions will be discussed:

1) Does the dashed line necessarily exist? Is V always associated with S following restriction to I = *yes*, as assumed?

2) Are there crucial missing arrows?

3) Are the vaccinated always more likely to be highly susceptible than the unvaccinated (positive association), as the authors assume? Again, the direction determines the direction of the bias.

It will be shown that the answer to the first and third questions is "No", and the answer to the second question is "Yes".

4. Discussion

There is a theorem on the topic of the first question [9]: V is associated with S when I = *yes* if and only if they modify each other's effects, on a risk (probability) ratio scale, on I = *yes*. Otherwise, V and S remain unassociated. Simple algebra can show that effect modification on the risk ratio scale is equivalent to interaction on the multiplicative scale.

If effect modification is weak (weak interaction), vaccine effectiveness (1 minus the risk ratio) does not vary much by susceptibility. Early on, it was reported that the mRNA vaccines were highly effective against infection in various subpopulations, which likely reflect different levels and varying distributions of susceptibil-

ity (e.g., age groups). Insignificant effect modification implies insignificant colliding bias. The dashed line might not exist at all, or it might be an insignificant source of bias.

If the dashed line exists, what is the direction of the induced association?

The authors write [4]:

"*An individual who becomes infected despite receiving a vaccine dose is more likely to have high susceptibility to infection than an unvaccinated individual who becomes infected. Therefore, among infected people, a greater proportion of people with high susceptibility to infection would be expected among those who received the vaccine than among those who did not receive the vaccine…*" (italics added)

That's a positive association between V and S (**Figure 1**), which is the key point of both articles. Although the explanation sounds intuitive, the formal explanation is more complicated and reveals the underlying assumptions. It is connected to the interaction between V and S on I = *yes*.

Typical examples of interaction consider two variables that affect the outcome in the same direction, like one of the authors' examples [4]:

"*Among successful actors, being physically attractive is inversely related to being a good actor.*" (italics added)

Both physical attractiveness and good acting increase the probability of a successful acting career. That's a positive interaction, which implies a negative association between the two causes, given a successful acting career. Other examples in epidemiological literature consider two risk factors and a bad outcome (resulting, again, in a negative association).

The situation is not so simple when a protective factor (a vaccine) interacts with a risk factor (susceptibility). The difficulty can be avoided by reversing the values of the protective factor and considering an interaction between *not taking a vaccine* and susceptibility. Now both operate to increase the probability of an infection.

If there is a positive interaction, these variables will be negatively associated: Non-vaccination implies low susceptibility, and vaccination implies high susceptibility. Reverting to the natural coding of a protective factor, V and S are positively associated. A plus sign should be placed next to the dashed line (**Figure 1**). That's what the authors infer in the key quote above.

But what is the nature of a positive interaction between V and S from the perspective of effect modification?

A few lines of algebra provide an answer (below). P is the probability (risk) of infection, stratified on vaccination status and susceptibility status. The first subscript applies to V, and 1 denotes *non-vaccination*. The second subscript applies to susceptibility (1 = high).

A positive interaction means that the joint effect is greater than the product of the separate effects:

$$\frac{P_{11}}{P_{00}} > \frac{P_{10}}{P_{00}} \times \frac{P_{01}}{P_{00}}$$

We may rewrite the inequality as a modification of the non-vaccination effect (risk ratio > 1) by the level of susceptibility:

$$\frac{P_{11}}{P_{01}} > \frac{P_{10}}{P_{00}}$$

Taking the inverse, we get the effect of vaccination (risk ratio < 1) on infection by the level of susceptibility:

$$\frac{P_{01}}{P_{11}} < \frac{P_{00}}{P_{10}}$$

The effect of vaccination is stronger when susceptibility is high than when it is low. That's the implicit assumption behind the plus sign next to the dashed line.

And reciprocally:

The effect of susceptibility (*risk ratio* > 1) *is weaker in the vaccinated than in the unvaccinated*.

The last result sounds intuitively true when a vaccine is indeed protective. If so, we are inclined to conclude that effect modification between V and S likely exists.

That's the positive side of the inference, but it comes with a negative side. As we already realized, the effectiveness of the vaccine must be lower when susceptibility is low. Is this indeed the case?

Again, that was not the message about Covid vaccines. As we recall, they were reported to have had over 90% effectiveness against infection (risk ratios < 0.1) in numerous subpopulations. Moreover, if effect modification was present, it seemed to have been weak and in the opposite direction. For example,

"*We estimated vaccine effectiveness in relation to various numbers of coexisting conditions and found indications that effectiveness may be slightly lower among persons with higher numbers of coexisting conditions*." [10] (Italics added)

Presumably, those coexisting conditions are part of the susceptibility construct.

Most importantly, one component of susceptibility is the function of the immune system. People with compromised immune function are susceptible to infection because exposure to the virus does not generate a strong immune response to its antigens. Nor do they respond well to the mRNA-generated spike protein [11]. Many of them fail to produce high titers of anti-SARS-CoV-2 antibodies post vaccination [12].

Which takes us to the second question: Are there crucial missing arrows in **Figure 1**?

Yes, there are, and they should be separated out from other domains of stable, pre-baseline susceptibility (**Figure 2**).

The vaccine operates to present the antigen before infection in order to generate an immune response (antibodies) that will prevent infection. The purpose of vaccination is to reduce this component of the construct called "level of susceptibility".

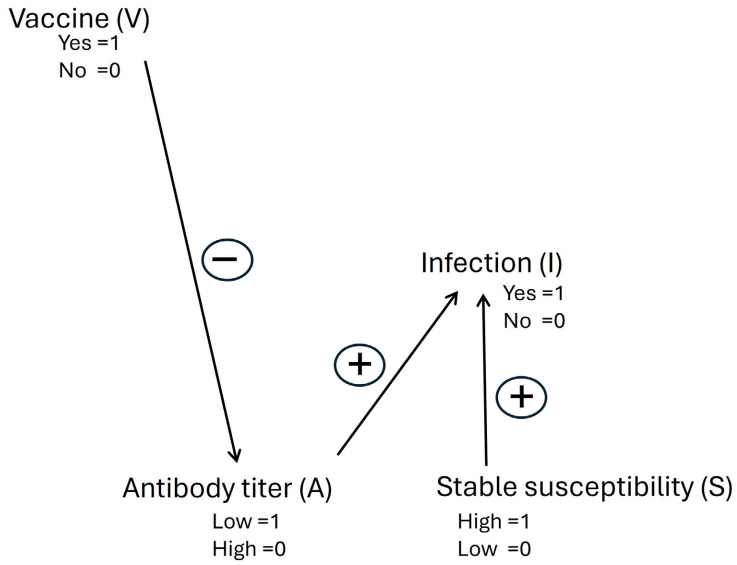

Figure 2. Vaccination prevents infection by increasing antibody titers.

5. An Alternative Causal Structure

The authors proposed one diagram. I propose another (Figure 3). It is far more complicated, but the focus is on two partially overlapping paths: V → A → J and V---A → J.

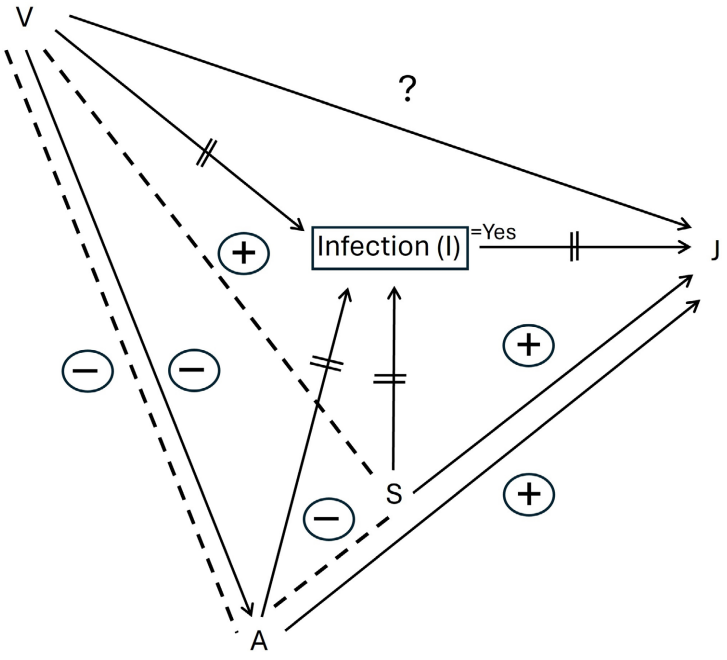

Figure 3. An alternative causal diagram on the assumption of effect modification (interaction) between any two causes of infection.

The sign next to a dashed line indicates the assumed direction. The diagram contains the key elements of the authors' assumption: a positive association between vaccination and stable susceptibility.

Immune imprinting is not an all-or-none phenomenon. Exposure to the previous antigen does not shut down every mechanism of protection against a new variant. The (protective) path V → A → J still exists. For example:

"Some neutralizing antibodies can target epitopes within the stem or stalk regions, which are relatively conserved between viral variants, are less susceptible to immune escape, and can be broadly protective. Non-neutralizing antibodies that target subdominant epitopes or regions in the stem or stalk play an important role in recruiting innate and cellular antiviral immunity." [5] (Italics added)

The underlying causal reality is not fully captured by any causal diagram, but **Figure 3** is more complete than **Figure 1**. The vaccinated are less likely to have a second infection or other infection-related outcomes (J = *yes*) through the path V → A → J, opposite to the path of colliding bias (V---S → J). We do not know exactly when the causal path is completely nullified because of waning immunity.

The explanation of the path V---A → J is more complicated.

The pair {V, A} affects I = *yes* similarly to the pair {V,S}. Recall, however, that a positive association between V and S when I = *yes* implied a stronger effect of vaccination when susceptibility was high (S = 1). Do we expect a stronger effect of vaccination when A = 1 (low titer)? People with low baseline antibody titers include those with compromised immunity. Empirical evidence indicates that they don't respond as well to a Covid vaccine, and vaccine effectiveness might be lower [11]. The effect of V on I might be *weaker* when A = 1, and the algebraic trail leads back to a *negative* association between V and A after conditioning, which amplifies the negative association along the causal arrow (**Figure 3**).

Which takes us to the third question: Are the vaccinated always more likely to be highly susceptible than the unvaccinated (positive association)?

No, we observe an induced path (V---A) where the vaccinated are less likely to be susceptible than the unvaccinated, and the direction of the "sum" of natural and induced paths between V and J is unpredictable. Underestimation of the true effect of immune imprinting is equally possible.

6. Conclusions

The generic lesson is not new. Methodology can sometimes tell consequences under different causal assumptions. It cannot tell which set of assumptions is true. If the purpose was to show how colliding bias might be added, that's a helpful contribution. If the purpose was to discredit observational studies of immune imprinting [3] [4], the authors went too far in assuming the direction and magnitude of the bias.

Knowledge is advanced when an alternative theoretical explanation turns into a testable hypothesis [13]. Modification of the vaccination effect by "an unmeasured susceptibility variable", just like "confounding from incomplete adjustment", is possible—but truism does not add knowledge. In contrast, modification of the effect by a coexisting condition, or by a state of impaired immunity, can be tested and has been tested to some extent. If colliding bias is significant at all, so far it

seems to operate in the opposite direction.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Shahar, E. and Shahar, D.J. (2012) Causal Diagrams and Three Pairs of Biases. In: Lunet, N., Ed., *Epidemiology-Current Perspectives on Research and Practice*, IntechOpen, 31-62. https://doi.org/10.5772/33486

[2] Shahar, E. (2025) "Selection Bias Without Colliders" Was… Effect Modification Bias x 3. In: Shahar, E., Ed., *Topics in Epidemiology and Statistics*. Amazon Kindle E-Books, 524. https://www.amazon.com/Topics-Epidemiology-Statistics-students-teachers-ebook/dp/B0DT78N6VY

[3] Monge, S., Pastor-Barriuso, R. and Hernán, M.A. (2023) The Imprinting Effect of Covid-19 Vaccines: An Expected Selection Bias in Observational Studies. *BMJ*, **381**, e074404. https://doi.org/10.1136/bmj-2022-074404

[4] Hernán, M.A. and Monge, S. (2023) Selection Bias Due to Conditioning on a Collider. *BMJ*, **381**, p1135. https://doi.org/10.1136/bmj.p1135

[5] Maltseva, M., Keeshan, A., Cooper, C. and Langlois, M. (2024) Immune Imprinting: The Persisting Influence of the First Antigenic Encounter with Rapidly Evolving Viruses. *Human Vaccines & Immunotherapeutics*, **20**, Article 2384192. https://doi.org/10.1080/21645515.2024.2384192

[6] Greenland, S., Pearl, J. and Robins, J.M. (1999) Causal Diagrams for Epidemiologic Research. *Epidemiology*, **10**, 37-48. https://doi.org/10.1097/00001648-199901000-00008 https://journals.lww.com/epidem/Abstract/1999/01000/Causal_Diagrams_for_Epidemiologic_Research.8.aspx

[7] Jiang, Z. and Ding, P. (2017) The Directions of Selection Bias. *Statistics & Probability Letters*, **125**, 104-109. https://doi.org/10.1016/j.spl.2017.01.022

[8] Nguyen, T.Q., Dafoe, A. and Ogburn, E.L. (2019) The Magnitude and Direction of Collider Bias for Binary Variables. *Epidemiologic Methods*, **8**, Article 20170013. https://doi.org/10.1515/em-2017-0013

[9] Shahar, D.J. and Shahar, E. (2017) A Theorem at the Core of Colliding Bias. *The International Journal of Biostatistics*, **13**, Article 20160055. https://doi.org/10.1515/ijb-2016-0055

[10] Dagan, N., Barda, N., Kepten, E., Miron, O., Perchik, S., Katz, M.A., *et al.* (2021) Bnt162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting. *New England Journal of Medicine*, **384**, 1412-1423. https://doi.org/10.1056/nejmoa2101765

[11] Napuri, N.I., Curcio, D., Swerdlow, D.L. and Srivastava, A. (2022) Immune Response to COVID-19 and mRNA Vaccination in Immunocompromised Individuals: A Narrative Review. *Infectious Diseases and Therapy*, **11**, 1391-1414. https://doi.org/10.1007/s40121-022-00648-2

[12] Barnes, E., Goodyear, C.S., Willicombe, M., Gaskell, C., Siebert, S., I de Silva, T., *et al.* (2023) SARS-CoV-2-Specific Immune Responses and Clinical Outcomes after COVID-19 Vaccination in Patients with Immune-Suppressive Disease. *Nature Medicine*, **29**, 1760-1774. https://doi.org/10.1038/s41591-023-02414-4

[13] Shahar, E. and Shahar, D.J. (2010) More on Selection Bias. *Epidemiology*, **21**, 429-430. https://doi.org/10.1097/ede.0b013e3181d7ec13
https://www.jstor.org/stable/25662886