Double Negative Control Inference in Test-Negative Design Studies of Vaccine Effectiveness

Kendrick Qijun Li*
Department of Biostatistics, University of Michigan;
Xu Shi
Department of Biostatistics, University of Michigan;
Wang Miao
Department of Probability and Statistics, Peking University;
Eric Tchetgen Tchetgen
Department of Statistics and Data Science, The Wharton School,
University of Pennsylvania

March 29, 2022

Abstract

The test-negative design (TND) has become a standard approach to evaluate vaccine effectiveness against the risk of acquiring infectious diseases in real-world settings, such as Influenza, Rotavirus, Dengue fever, and more recently COVID-19. In a TND study, individuals who experience symptoms and seek care are recruited and tested for the infectious disease which defines cases and controls. Despite TND's potential to reduce unobserved differences in healthcare seeking behavior (HSB) between vaccinated and unvaccinated subjects, it remains subject to various potential biases. First, residual confounding bias may remain due to unobserved HSB, occupation as healthcare worker, or previous infection history. Second, because selection into the TND sample is a common consequence of infection and HSB, collider stratification bias may exist when conditioning the analysis on testing, which further induces confounding by latent HSB. In this paper, we present a novel approach to identify and estimate vaccine effectiveness in the target population by carefully leveraging a pair of negative control exposure and outcome variables to account for potential hidden bias in TND studies. We illustrate our proposed method with extensive simulation and an application to study COVID-19 vaccine effectiveness using data from the University of Michigan Health System.

Keywords: Causal inference, proximal causal inference, selection bias, unmeasured confounding

Contents

1 Introduction

1.1 Text-negative design studies of vaccine effectiveness

The test-negative design (TND) has become a standard approach to evaluate real-world vaccine effectiveness (VE) against the risk of acquiring infections diseases (Chung et al., 2020; Flannery et al., 2019; Jackson et al., 2017; Rolfes et al., 2019; Tenforde et al., 2021). In an outpatient Influenza VE test-negative design, for example, symptomatic individuals seeking care and meeting eligibility criteria are enrolled and their Influenza virus infection status is subsequently confirmed via a laboratory test. VE against flu infection is then measured by comparing the prevalence of vaccination between the test-positive "cases" and test-negative "controls" (Jackson and Nelson, 2013; Jackson et al., 2017). Besides Influenza, the TND and its variants have also been applied to study VE against pneumococcal disease (Broome, Facklam, and Fraser, 1980), dengue (Anders et al., 2018; Utarini et al., 2021), rotavirus (Boom et al., 2010; Schwartz et al., 2017), and other infectious diseases. Recently, the TND has increasingly been used in post-licensure evaluation of COVID-19 VE (Dagan et al., 2021; Dean, Hogan, and Schnitzer, 2021; Hitchings et al., 2021; Olson et al., 2022; Patel, Jackson, and Ferdinands, 2020; Thompson et al., 2021).

Test-negative designs are believed to reduce unmeasured confounding bias due to health-care seeking behavior (HSB), whereby care seekers are more likely to be vaccinated, have healthier behaviors that reduces the risk of infection, and get tested when ill (Jackson et al., 2006; Shrank, Patrick, and Brookhart, 2011). By restricting analysis to care seekers who are tested for the infection in view (e.g. Influenza or COVID-19), the vaccinated and unvaccinated are more likely to share similar HSB and underlying health characteristics. Misclassification of infection status is also reduced because the analysis is restricted to

tested individuals (Jackson and Nelson, 2013).

Sullivan et al. (2016) used directed acyclic graphs (DAG) to illustrate the rationale behind TND in the context of evaluating VE against Influenza infection, as shown in Figures 1(a) and (b). We denote Influenza vaccination status by A and Influenza infection by Y, so that the arrow $A \rightarrow Y$ represents VE against flu infection. Selection into the TND study sample, denoted by S, is triggered by a subject experiencing flu-like symptoms or acute respiratory illness, seeking care at clinics or hospitals, and getting tested for Influenza infection, hence the $Y \rightarrow S$ edge. Healthcare seeking behavior, denoted by HSB, may affect S, A, and Y because subjects with certain health seeking proclivities may be more likely to seek care, take annual flu shots, and participate in healthy and preventative behaviors. The above variables are subject to effects of other clinical or demographic factors, such as age, season and high-risk conditions, included in Figure 1 as confounders X. The TND assumes that by restricting recruitment to care seekers, the study subjects have identical healthcare seeking behavior; in other words, conditioning the analysis on $S = 1$ necessarily leads to HSB$= 1$, which blocks the effects of HSB (Figure 1(b)). The effects of X are further adjusted for by including these factors in a logistic regression model or by inverse probability weighting (Bond, Sullivan, and Cowling, 2016; Thompson et al., 2021).

However, the TND remains subject to potential hidden bias. First, the assumption that all study subjects seeking care are lumped into a single category HSB$= 1$ may be unrealistic. It may be more realistic that HSB is not a deterministic function of S and remains a source of confounding bias even after conditioning on S. Furthermore, there might be other mismeasured or unmeasured confounders, denoted as U. For example, healthcare workers are at increased risk of flu infection due to higher exposure to flu patients and are more likely to seek care and receive vaccination due to health agency guide-

lines (Black et al., 2018). Previous flu infection history may also be a source of confounding if it alters the likelihood of vaccination and care seeking, while also providing immunity against circulating strains (Krammer, 2019; Sullivan, Tchetgen Tchetgen, and Cowling, 2016). These unmeasured or mismeasured potential sources of confounding, if not properly accounted for, can result in additional confounding bias, as illustrated in Figure 1(c). Finally, collider stratification bias is likely present due to conditioning on S, which is a common consequence of HSB, other risk factors (X, U), and Influenza infection Y (Lipsitch, Jha, and Simonsen, 2016). That is, conditioning on S unblocks the backdoor path $A \leftarrow (X, U, HSB) \rightarrow S \leftarrow Y$, which would in principle be blocked if study subjects had identical levels of HSB and other risk factors (Sullivan, Tchetgen Tchetgen, and Cowling, 2016).

Accounting for these potential sources of bias is well known to be challenging, and potentially infeasible without additional assumptions or data. This can be seen in Figure 1(d), which is a simplified version of Figure 1(c) where the unmeasured confounders U include individuals' occupation as a healthcare worker, previous flu infection, HSB, and so on. Figure 1(d) indicates that the unmeasured confounders U induce both confounding bias through the path $A \leftarrow U \rightarrow Y$ and collider stratification bias through the path $A \leftarrow U \rightarrow S \leftarrow Y$. In presence of both unmeasured confounding and collider bias, causal bounds may be available (Gabriel, Sachs, and Sjölander, 2020) but likely too wide to be informative; causal identification in TND therefore remains to date an important and outstanding open problem in the causal inference literature which we aim to resolve.
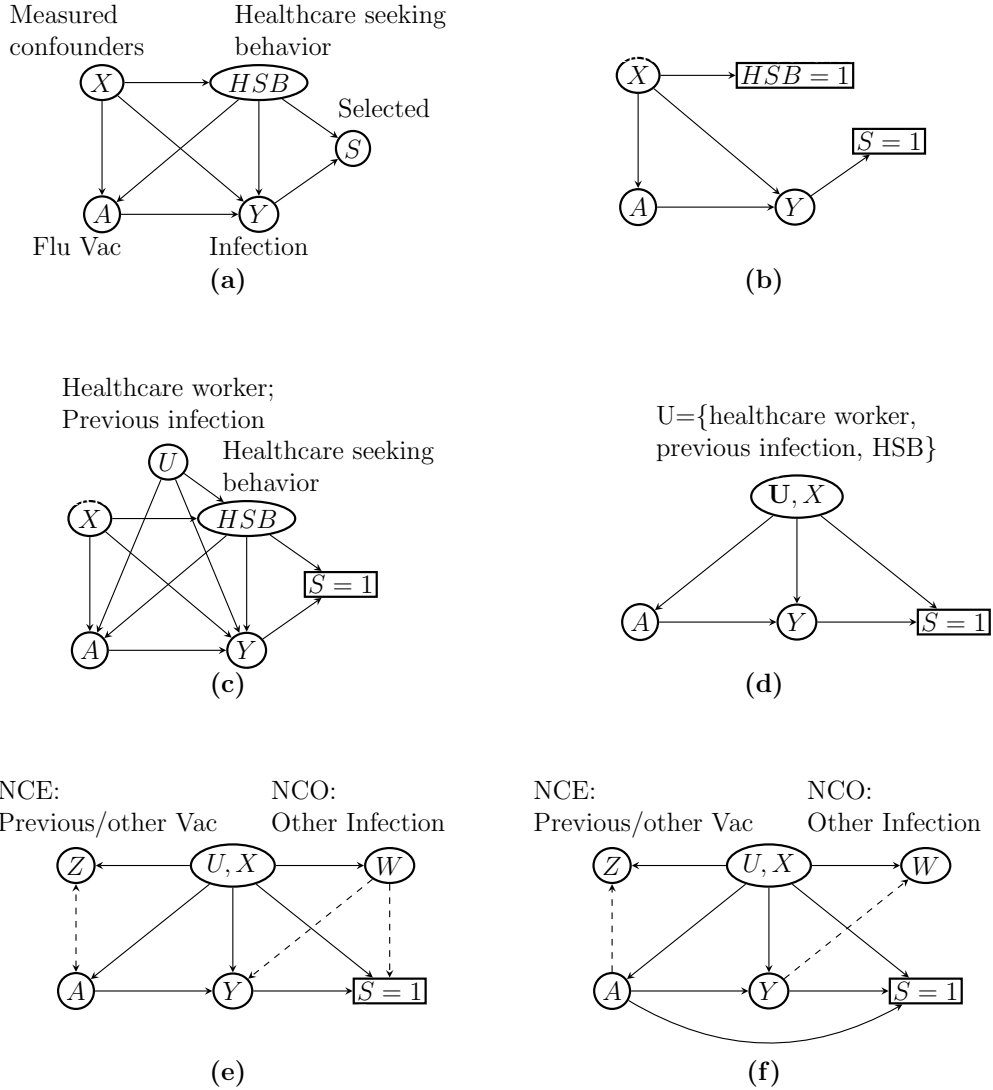
Figure 1: Causal relationships of variables in a test-negative design. Sullivan, Tchetgen Tchetgen, and Cowling, 2016 used (a) to illustrate the causal relationship between variables in a test-negative design in the general population, and used (b) to illustrate the assumption implicit in the common approach to estimate VE from the study data that study subjects have identical healthcare seeking behavior (HSB) (Sullivan, Tchetgen Tchetgen, and Cowling, 2016). (c) shows that if HSB remains partially unobserved, then the backdoor paths $A \leftarrow HSB \rightarrow Y$ and $A \leftarrow HSB \rightarrow S = 1 \leftarrow Y$ indicate unmeasured confounding bias and selection bias, respectively. Other unmeasured confounders, such as occupation as a healthcare worker and previous infection, open additional backdoor paths between A and Y and result in additional confounding bias. (d) shows a simplified DAG from (c) that combines the unmeasured confounders into a single variable U. (e) illustrates our approach to estimate VE leveraging negative control exposure Z and outcome W. Dashed arrows indicate effects that are not required. (f) shows a scenario with the $A \rightarrow Y$ arrow where the causal odds ratio can still be identified under additional assumptions.

8

1.2 Negative control methods

In recent years, negative control variables have emerged as powerful tools to detect, reduce and potentially correct for unmeasured confounding bias (Lipsitch, Tchetgen Tchetgen, and Cohen, 2010; Miao, Geng, and Tchetgen Tchetgen, 2018; Shi, Miao, and Tchetgen Tchetgen, 2020). The framework requires that at least one of two types of negative control variables are available which are *a priori* known to satisfy certain conditions: a negative control exposure (NCE) known to have no direct effect on the primary outcome; or a negative control outcome (NCO), known not to be an effect of the primary exposure. Such negative control variables are only valid and therefore useful to address unmeasured confounding in a given setting to the extent that they are subject to the same source of confounding as the exposure-outcome relationship of primary interest. Thus, the observed association between a valid NCE and the primary outcome (conditional on the primary treatment and observed covariates) or that between a valid NCO and the primary exposure can indicate the presence of residual confounding bias. For example, in a cohort study to investigate flu VE against hospitalization and death among seniors, to detect the presence of confounding bias due to underlying health characteristics, Jackson et al. (2006) used hospitalization/death before and after the flu season as NCOs and found that the association between flu vaccination and hospitalization was virtually the same before and during the flu season, suggesting that the lower hospitalization rate observed among vaccinated seniors versus unvaccinated seniors was partially due to healthy user bias.

Recently, new causal methods have been developed to not only detect residual confounding when present, but also to potentially de-bias an observational estimate of a treatment causal effect in the presence of unmeasured confounders when both an NCE and an NCO are available, referred to as the double negative control (Miao, Geng, and Tchetgen Tch-

etgen, 2018; Shi et al., 2020; Tchetgen Tchetgen et al., 2020). In this recent body of work, the double negative control design was extended in several important directions including settings in which proxies of treatment and outcome confounding routinely measured in well designed observational studies may be used as negative control variables, a framework termed *proximal causal inference*; longitudinal settings where one is interested in the joint effects of time-varying exposures (Ying et al., 2021), potentially subject to both measured and unmeasured confounding by time-varying factors; and in settings where one aims to estimate direct and indirect effects in mediation analysis subject to unmeasured confounding or unmeasured mediators (Dukes, Shpitser, and Tchetgen Tchetgen, 2021; Ghassami, Shpitser, and Tchetgen Tchetgen, 2021). Additional recent papers in this fast-growing literature include Qi, Miao, and Zhang (2021), Liu and Tchetgen Tchetgen (2021), Egami and Tchetgen Tchetgen (2021), Kallus, Mao, and Uehara (2021), Imbens, Kallus, and Mao (2021), Deaner (2018), Deaner (2021), Ghassami et al. (2021), Mastouri et al. (2021) and Ghassami, Shpitser, and Tchetgen Tchetgen (2022). Importantly, existing identification results in negative control and proximal causal inference literature has been restricted to i.i.d settings (Miao, Shi, and Tchetgen Tchetgen, 2018) and time series settings (Shi et al., 2021), and to date, to the best of our knowledge, outcome-dependent sampling settings such as TND, have not been considered, particularly one where confounding and selection bias might co-exist.

1.3 Outline

The rest of the paper is organized as followed: we introduce notation and the identification challenge in view in Sections 2.1. Next we develop the identification strategy and describe a new debiased estimator under a double negative control TND study in Section 2.2-2.4,

assuming (1) homogeneous VE across strata defined by all measured and unmeasured confounders and (2) no direct effect of vaccination on selection into the TND sample. In Section 2.5, we relax the homogeneous VE assumption and describe identification and estimation allowing for VE to depend on observed covariates. In Section 2.6, we relax the assumption of no direct effect of vaccination on selection and introduce the assumptions under which our VE estimator is unbiased on the odds ratio scale. In Section 3, we demonstrate the performance of our method with simulation. In Section 4, the approach is further illustrated in an application to estimate COVID-19 VE against infection in a TND study nested within electronic health records from University of Michigan Health System. We then conclude with a discussion in Section 5.

2 Method

2.1 Preliminary: estimation under no unmeasured confounding and no selection bias

To fix ideas, we first review estimation assuming all confounders (U, X) are fully observed and the study sample is randomly drawn (rather than selected by testing) from source population, referred to as the "target population". That is, we observe data on (A, Y, U, X) which are independent and identically distributed in the target population. For each individual, we write $Y(a)$ as the binary potential infection outcome had, possibly contrary to fact, the person's vaccination status been $A = a$, $a = 0, 1$. Our goal is to provide identification and estimation strategies for the causal risk ratio (RR) defined as $RR = E[Y(1)]/E[Y(0)]$. Let β_0 denote the log causal RR, i.e., $RR = \exp(\beta_0)$. Following Hudgens and Halloran (2006) and Struchiner and Halloran (2007), we define VE as one minus the causal RR:

$VE = 1 - RR = 1 - \exp(\beta_0)$. The potential outcomes and the observed data are related through the following assumptions:

Assumption 1. *(Identification conditions of mean potential outcomes).*

(a) *(Consistency)* $Y(a) = Y$ *if* $A = a$ *almost surely for* $a = 0, 1$;

(b) *(Exchangeability)* $A \perp\!\!\!\perp Y(a)|U, X$ *for* $a = 0, 1$.

(c) *(Positivity)* $0 < P(A = a|U, X)$ *almost surely* $a = 0, 1$.

Assumption 1(a) states that the infection status of a subject with vaccination status $A = a$ is equal to the corresponding potential outcome $Y(a)$. This further requires that the treatment is sufficiently well-defined and a subject's potential outcome is not affected by the treatment of other subjects (Cole and Frangakis, 2009). Assumption 1(b) states that treatment is exchangeable within strata of (U, X), i.e. there is no unmeasured confounding given (U, X). We develop methods that allow U to be unmeasured in Section 2.3. Assumption 1(c) states that for all realized values of (U, X) there is at least one individual with an opportunity to get treatment $a = 0, 1$.

Let $Q(A = a, U, X) = 1/P(A = a|U, X)$ denote the inverse of the probability of vaccination status $A = a$ given confounders (Rosenbaum, 1987). Under Assumption 1, it is well known that, if U were observed, the mean potential outcome for treatment a in the general population can be identified by inverse probability of treatment weighting (IPTW):

$$E[Y(a)] = E\left[I(A = a)Q(A = a, U, X)Y\right], \tag{1}$$

for $a = 0, 1$. Therefore, the log causal RR β_0 satisfies the following equation

$$E\left[Q(A = 1, U, X)AY \exp(-\beta_0)\right] - E\left[Q(A = 0, U, X)(1 - A)Y\right] = 0.$$

Equivalently, we have

$$E[V_0(A, Y, U, X; \beta_0)] = 0 \qquad (2)$$

for unbiased estimation, where $V_0(A, Y, U, X; \beta) = (-1)^{1-A}Q(A, U, X)Y \exp(-\beta A)$.

2.2 Tackling selection bias under a semiparametric risk model

Next, consider a TND study for which data (A, Y, X, U) is observed only for the tested individuals with $S = 1$. Because S is impacted by other factors such as infection, the estimating function $V_0(A, Y, U, X; \beta_0)$ may not be unbiased with respect to the study sample; i.e. $V_0(A, Y, U, X; \beta_0)|U, X, S = 1] \neq 0$ without another assumption about the selection process into the TND sample.

For a study sample of size n from a TND, we denote the i-th study subject's variables as (A_i, Y_i, U_i, X_i), $i = 1, \ldots, n$. For generalizability, we first make the key assumption that vaccination A is unrelated to selection S other than through a subject's infection status Y and confounders (U, X).

Assumption 2 (Treatment-independent sampling). $S \perp\!\!\!\perp A|Y, U, X$.

In the test-negative design, this assumption requires that an individual's decision to seek care and get tested only depends on the presence of symptoms and his/her underlying behavioral or socioeconomic characteristics, including HSB (contained in (U, X)); a person's vaccination status does not directly Influence their selection process. The DAGs in Figure 1(a)-(e) in fact encode this conditional independence condition. We will relax this assumption in Section 2.6.

Assumption 3 (No effect modification by a latent confounder). *For $a = 0, 1$,*

$$P(Y = 1|A = a, U, X) = \exp(\beta_0 a)g(U, X) \qquad (3)$$

13

where $g(U, X)$ is an unknown function only restricted by $0 \leq P(Y = 1|A, U, X) \leq 1$.

Assumption 3 defines a semiparametric multiplicative risk model which posits that vaccine effectiveness, measured on the RR scale, is constant across (U, X) strata in the target population. In other words, the effect of vaccination A on the risk of infection Y is not modified by confounders U, X. In Section 2.5, we will relax the assumption to allow for effect modification by measured confounders X. Infection risk for control subjects $P(Y = 1|A = 0, U, X) = g(U, X)$ is left unspecified and thus defines the nonparametric component of the model.

Under Assumptions 1 and 3, one can verify that $\exp(\beta_0) = E[Y(1)]/E[Y(0)]$, which is the marginal causal RR. The potential infection outcome means $E[Y(1)]$ and $E[Y(0)]$ in the target population cannot be identified due to the study selection process without an additional restriction. Nevertheless, the estimating equation (2) implies that it may still be possible to identify β_0 without necessarily identifying $E[Y(0)]$ and $E[Y(1)]$. The following proposition indicates that the same is true when the data are subject to selection bias of certain structure.

Proposition 1. *Under Assumptions 1- 3, the parameter β_0 satisfies*

$$E[V_0(A, Y, U, X; \beta_0) \mid U, X, S = 1] = 0. \tag{4}$$

The proof of Proposition 1 is in Appendix A. From Proposition 1, the IPTW estimating function V_0 derived from the target population is also unbiased with respect to the study sample.

Under Assumptions 1-3, one can estimate β_0 with $\hat{\beta}$ as the solution to

$$\frac{1}{n}\sum_{i=1}^{n}(-1)^{1-A_i}c(X_i)\widehat{Q}(A_i, U_i, X_i)Y_i \exp(-\hat{\beta}_0 A_i) = 0, \tag{5}$$

where n is the size of the selected sample, $c(\cdot)$ is a user specified function, and $\widehat{Q}(A_i, U_i, X_i) = \widehat{P}(A = A_i \mid U_i, X_i)$ is the estimated probability of have vaccination status $A = A_i$ given confounders (U_i, X_i). Letting $c(X_i) = 1$, the resulting estimator

$$\hat{\beta}_0 = \left[\sum_{i=1}^{n}\hat{Q}(A_i, U_i, X_i)A_i Y_i\right] / \left[\sum_{i=1}^{n}\hat{Q}(A_i, U_i, X_i)(1 - A_i)Y_i\right]$$

is essentially the IPTW estimator of marginal RR in Schnitzer (2022) assuming (U_i, X_i)'s are all observed.

However, $Q(A, U, X)$ cannot be estimated because U is unobserved. Furthermore, even if U were observed, $\widehat{Q}(A_i, U_i, X_i)$ may not be identified from the TND sample due to selection bias. In the next section, we describe a new framework to account for unmeasured confounding in a TND setting, leveraging negative control exposure and outcome variables.

2.3 Tackling unmeasured confounding bias leveraging negative controls

2.3.1 Negative control exposure (NCE) and treatment confounding bridge function

As shown in Figure 1(e), suppose that one has observed a valid possibly vector-valued NCE, denoted as Z, which is a priori known to satisfy the following key independence conditions:

Assumption 4 (NCE independence conditions). $Z \perp\!\!\!\perp (Y, S)|A, U, X$.

Assumption 4 essentially states that any existing $Z - Y$ association conditional on (X, A) in the target population must be a consequence of their respective association with U, therefore indicating the presence of confounding bias. Importantly, the NCE must a priori be known to have no causal effect on infection status (Miao, Shi, and Tchetgen Tchetgen, 2018). Likewise, the association between Z and S conditional on (X, A) is completely due to their respective association with U. Figure 1(e) presents a graphical illustration of an NCE that satisfies Assumption 4.

In the Influenza VE setting, a candidate NCE can be vaccination status for the preceding year, or other vaccination status such as Tdap (Tetanus, Diphtheria, Pertussis) vaccine, as both are known to effectively provide no protection against the circulating flu strain in a given year. We now provide an intuitive description of our approach to leverage Z as an imperfect proxy of U for identification despite being unable to directly observe U.

To illustrate the rationale behind identification, ignore selection bias for now and suppose that $Q(A, U) = \alpha_0 + \alpha_1 A + \alpha_2 U$, suppressing measured confounders X. Although U is unobserved, suppose further that Z satisfies $E[Z|A, U] = \gamma_0 + \gamma_1 A + \gamma_2 U$. Then we have that

$$Q(A, U) = E[q(A, Z)|A, U], \qquad U = E[\tilde{U}(A, Z)|A, U],$$

where $\tilde{U} = (Z - \gamma_0 - \gamma_1 A)/\gamma_2$. Replacing U with $\tilde{U}$ in $Q(A, U)$, we get $q(A, Z) = \alpha_0 + \alpha_1 A + \alpha_2 \tilde{U}(A, Z)$, which does not depend on unmeasured confounder U and can represent the inverse probability of vaccination as $Q(A, U) = E[q(A, Z)|A, U]$. If all parameters of q were known, it would naturally follow that the IPTW method in (1) can be recovered by

$$E[Y(a)] = E\{I(A = a)E[q(A, Z)|A, U]Y\} \stackrel{A.4}{=} E[I(A = a)q(A, Z)Y],$$

Therefore, β_0 can be identified if the distribution of (A, Y, Z) in the target population is available provided that parameters indexing q can be identified.

The above insight motivates the following assumption:

Assumption 5 (treatment confounding bridge function). *There exists a function $q(A, Z, X)$ that satisfies, for every a, u and x,*

$$Q(A = a, U = u, X = x) = E\left[q(A, Z, X)|A = a, U = u, X = x\right] \tag{6}$$

The function q that satisfies (6) is called a treatment confounding bridge function, as it bridges the observed NCE with the unobserved propensity score (Cui et al., 2020). Below we give two examples where the integral equation (6) can easily be solved and the treatment confounding bridge function q admits a closed form solution.

Example 1. *(Binary U and Z) Suppose that U is binary, and so is the NCE Z. For simplicity we suppress X. The integral equation (6) can then be written as $\sum_{z=0}^{1} q(a, z)P(Z = z|U = u, A = a) = P(A = a|U = u)^{-1}$,*

or equivalently, $\sum_{z=0}^{1} p_{za.u}q(a, z) = 1$ for each $a, u \in \{0, 1\}$, where $p_{za.u} = P(Z = z, A = a|U = u)$. Therefore, the treatment confounding bridge function $q(a, z)$ solves the linear equation system

$$P_{Z,A|U}\begin{pmatrix} q(a,0) \\ q(a,1) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ where } P_{Z,A|U} = \begin{pmatrix} p_{0a.0} & p_{1a.0} \\ p_{0a.1} & p_{1a.1} \end{pmatrix}.$$

If the matrix $P_{Z,A|U}$ is invertible, then $q(a, z)$ has a closed form solution given by

$$q(a, z) = [p_{1a.1} - p_{1a.0} + (p_{0a.0} - p_{0a.1} - p_{1a.1} + p_{1a.0})z] / (p_{0a.0}p_{1a.1} - p_{0a.1}p_{1a.0}). \tag{7}$$

17

The result can be extended to the cases where Z is polytomous as detailed in Appendix D.

Example 2. *(Continuous U and Z) Suppose the unmeasured confounder U and the NCE Z are continuous. Further assume that*

$$A|U, X \sim Bernoulli([1 + \exp(-\mu_{0A} - \mu_{UA}U - \mu_{XA}X)]^{-1})$$

$$Z|A, U, X \sim N(\mu_{0Z} + \mu_{AZ}A + \mu_{UZ}U + \mu_{XZ}X, \sigma_Z^2).$$

By the derivation in Appendix E, the treatment confounding bridge function $q(A, Z, X)$ is

$$q(A, Z, X) = 1 + \exp\left[(-1)^A(\tau_0 + \tau_1 A + \tau_2 Z + \tau_3 X)\right] \tag{8}$$

where $\tau_0 = \mu_{0A} - \dfrac{\mu_{UA}\mu_{0Z}}{\mu_{UZ}} - \dfrac{\sigma_z^2 \mu_{UA}^2}{2\mu_{UZ}^2}$, $\tau_1 = \dfrac{\sigma_z^2 \mu_{UA}^2}{\mu_{UZ}^2} - \dfrac{\mu_{UA}\mu_{AZ}}{\mu_{UZ}}$, $\tau_2 = \mu_{UA}/\mu_{UZ}$, and $\tau_3 = \mu_{XA} - \mu_{XZ}\mu_{UA}/\mu_{UZ}$.

Formally, Equation (6) defines a Fredholm integral equation of the first kind, with treatment confounding bridge function $q(A, Z, X)$ as its solution (Cui et al., 2020). Heuristically, the existence of a treatment confounding bridge function requires that variation in Z induced by U is sufficiently correlated with variation in A induced by U. For instance, in Example 1, existence of a treatment confounding bridge function requires that the matrix $P_{Z,A|U}$

is nonsingular. In Example 2, the existence of a treatment confounding bridge function amounts to the condition $\mu_{UZ} \neq 0$, which again requires $Z \not\perp\!\!\!\perp U|A, X$. Cui et al. (2020) provided formal conditions sufficient for the existence of the treatment confounding bridge function satisfying Equation (6). These conditions are reproduced for completeness in Appendix B.

Thus, under Assumption 5, we propose to construct a new unbiased estimating function

18

for β_0 by replacing $Q(A, U, X)$ with $q(A, Z, X)$ in $V_0(A, Y, U, X; \beta_0)$.

Theorem 1. *(Moment restriction of β_0) Under Assumptions 1-5, we have that*

$$E[V_1(A, Y, Z, X; \beta_0) \mid U, X, S = 1] = 0$$

where $V_1(A, Y, Z, X; \beta_0) = (-1)^{1-A} q(A, Z, X) Y \exp(-\beta_0 A)$.

The proof of Theorem 1 is in Appendix C. In practice, if one can consistently estimate the treatment confounding bridge function $q(A, Z, X)$ with $\widehat{q}(A, Z, X)$, Theorem 1 suggests estimating β_0 by solving the estimating equation

$$\frac{1}{n} \sum_{i=1}^{n} (-1)^{1-A_i} c(X_i) \widehat{q}(A_i, Z_i, X_i) Y_i \exp(-\beta_0 A_i) = 0, \tag{9}$$

which, for a one-dimensional $c(\cdot)$, results in a closed form estimator

$$\widehat{\beta}_0 = \log \left(\frac{\sum c(X_i) \widehat{q}(A_i, Z_i, X_i) A_i Y_i}{\sum c(X_i) \widehat{q}(A_i, Z_i, X_i)(1 - A_i) Y_i} \right).$$

Importantly, although (6) may not have a unique solution, any solution uniquely identifies the causal log RR β_0 for a fixed function $c(\cdot)$. Furthermore, although the choice of $c(\cdot)$ doesn't change the unbiasedness of (9), it impacts the validity and efficiency of the resulting estimator $\widehat{\beta}_0$. For example, (9) doesn't lead to a estimator of β_0 if $c(X_i)$ is set to be zero. In practice, one may simply set $c(X_i) \equiv 1$. The result in Theorem 1 cannot directly be applied in practice because the treatment confounding bridge function is not identifiable even if random samples from the target population were available – solving (6) requires additional information about U which is unobserved. For instance, in Example 1 one is unable to directly estimate $q(a, z)$ because $p_{za.u}$ in (7) cannot directly be estimated from

the observed data.

2.3.2 Negative control outcome (NCO) for identification of treatment confounding bridge function

For identification and estimation of q, we leverage negative control outcomes (NCO) to construct feasible estimating equations for the treatment confounding bridge function as in Cui et al. (2020). Similar to NCEs, NCOs can be viewed as imperfect proxies of U. However, unlike NCEs, a valid NCO, denoted by W, is a measured covariate which is (i) known a priori not to be a causal effect of either the primary exposure A nor the NCE Z; and (ii) is associated with (A, Z) conditional on X only to the extent that it is associated with U.

Formally, we make the following assumption.

Assumption 6. *(NCO Independence Conditions).*

(a) $W \perp\!\!\!\perp A|U, X$;   *(b)* $W \perp\!\!\!\perp Z|A, U, X, Y$;

(c) $S \perp\!\!\!\perp Z|A, U, X, W, Y$.

Assumptions 6(a) and (b) formalize the requirement that neither the primary exposure nor NCE have direct effects on the NCO. Assumption 6(c) complements Assumption 4 and states that conditioning on W in addition to (A, U, X, Y) does not alter the conditional independence of Z with S. In flu VE studies, a candidate NCO can be an infection whose risk is not causally affected by either A or Z. For example, if the NCE is selected to be Tdap vaccination, then a potential NCO may be current-year respiratory syncytial virus infection, as its risk is unlikely to be affected by Influenza or Tdap vaccination. Recent outpatient visits for other acute illnesses can also serve as NCO, such as blepharitis, wrist/hand sprain,

lipoma, ingrowing nail, etc. (Leung et al., 2011). Figure 1(e) illustrates an NCO W that satisfies Assumptions 6(a) and (b).

Similar to Cui et al. (2020), we leverage the availability of an NCO as an additional proxy to identify the treatment confounding bridge function. However, a complication arises due to lack of a random sample from the target population, a key requirement in the approach outlined in Cui et al. (2020). In general, it is not possible to obtain sufficient information about neither the distribution of W nor that of U in the target population from the TND data without an additional structural assumption (Bareinboim and Pearl, 2012). In the following, we avoid imposing such an additional structural assumption by leveraging an important feature of infectious diseases such as Influenza and COVID-19; mainly that contracting such an infection is a rare event in most target populations of interest, and therefore information from the target population that is relevant for estimating the treatment confounding bridge function can be recovered from the test-negative control group. Formally, we make the following rare disease assumption.

Assumption 7 (Rare infection)**.** *There exist a small positive number $\delta > 0$ such that*

$$P(Y = 1 | A = a, W = w, U = u, X = x) \leq \delta, \qquad for\ almost\ every\ a, w, u, x \qquad (10)$$

Assumption 7 states that infected subjects, whether vaccinated or not and regardless of their negative control outcomes, only constitute a small proportion of each (U, X) stratum in the general population; specifically, the assumption implies that $\frac{1}{1-\delta} \leq \frac{P(A,Z|U=u,X=x,Y=0)}{P(A,Z|U=u,X=x)} \leq 1 - \delta$. Thus, under Assumptions 2, 4 and 7, $P(A = a, Z = z | U = u, X = x) \approx P(A = a, Z = z | U = u, X = x, Y = 0, S = 1)$ for all a, z, x, u. We now introduce a key property of the treatment confounding bridge function in Theorem 2, which is proved in Appendix F.

Theorem 2 (Identification of the treatment confounding bridge function). *Under Assumptions 2, 4, 5, 6, and 7, for $a = 0, 1$ we have that*

$$\frac{(1-\delta)^3}{P(A = a|W, X, Y = 0, S = 1)} < E[q(a, Z, X)|W, A = a, X, Y = 0, S = 1]$$

$$< \frac{1}{(1-\delta)^3 P(A = a|W, X, Y = 0, S = 1)}$$

Thus, provided $\delta \approx 0$, Theorem 2 suggests that an approximation to the treatment confounding bridge function can be obtained by solving the following integral equation involving only observed data

$$E[q^*(A, Z, X)|W, A = a, X, Y = 0, S = 1] = 1/P(A = a|W, X, Y = 0, S = 1). \qquad (11)$$

as long as a solution exists. Accordingly, hereafter suppose that the following assumption holds.

Assumption 8 (Existence of a unique solution to (11)). *There exists a unique square-integrable function $q^*(A, Z, X)$ that satisfies (11).*

Heuristically, uniqueness of a solution to (11) requires that variation in W is sufficiently informative about variation in Z, in the sense that there is no source of variation in W that is not associated with a corresponding source of variation in Z. See Appendix G for further elaboration of completeness conditions and D'Haultfoeuille (2011) and Newey and Powell (2003) for related use of the assumption in the literature. Below we briefly illustrate Assumption 8 in the examples of Section 2.3.1.

Example 1′. *Suppose U and Z are both binary, and a binary NCO W is also observed. Let $p'_{za.w} = P(Z = z, A = a|W = w, Y = 0, S = 1)$ for $z, a, w \in \{0, 1\}$, then solving (11) is*

equivalent to solving the system of linear equations

$$p'_{0a.0}q^*(a,0) + p'_{1a.0}q^*(a,1) = 1; \qquad p'_{0a.1}q^*(a,0) + p'_{1a.1}q^*(a,1) = 1,$$

giving $q^*(a,z) = [p'_{1a.1} - p'_{1a.0} + (p'_{0a.0} - p'_{0a.1} - p'_{1a.1} + p'_{1a.0})z] / (p'_{0a.0}p'_{1a.1} - p'_{0a.1}p'_{1a.0})$. Note that the probabilities $p'_{za.w}$ can all be estimated from the study sample.

We emphasize that the solution to Equation (11) is ultimately an approximation to the (non-identifiable) treatment confounding bridge function in the target population. The accuracy of this approximation relies on the extent to which the rare disease assumption holds in the target population of interest. We study the potential bias resulting from a departure of this key assumption in the Appendix I. We further observe that, under the null hypothesis of no vaccine effectiveness, or if W has no direct effects on Y or S, then the function $q^*(A, Z, X)$ equals the treatment confounding bridge exactly, even for a non-rare disease outcome, as stated in the corollary below. We prove Corollary 1 in Appendix F.

Corollary 1. *Under the Assumptions of Theorem 1 and Assumption 8, if there is no vaccine effect against infection, such that $Y \perp\!\!\!\perp A|U, X$, then*

$$E[q(A, Z, X)|W, A = a, X, Y = 0, S = 1] = 1/P(A = a|W, X, Y = 0, S = 1).$$

From Theorem 2, we immediately have the following corollary which provides a basis for estimation of $q^*(A, Z, X)$ from the observed TND sample.

Corollary 2. *Under the conditions listed in Theorem 2, for any function $m(W, A, X)$, the solution $q^*(A, Z, X)$ to Equation (11) also solves the population moment equation*

$$E\left[m(W, A, X)q^*(A, Z, X) - m(W, 1, X) - m(W, 0, A)|Y = 0, S = 1\right] = 0. \qquad (12)$$

We prove Corollary 2 in Appendix H. In practical situations where a parametric model $q^*(A, Z, X; \tau)$ for the treatment confounding bridge function might be appropriate, where τ is an unknown finite dimensional parameter indexing the model, Corollary 2 suggests one can estimate τ by solving the estimating equation

$$\frac{1}{n}\sum_{i=1}^{n}(1 - Y_i)[m(W, A, X)q(A, Z, X; \tau) - m(W, 1, X) - m(W, 0, X)] = 0, \qquad (13)$$

where $m(W, A, X)$ is a user-specified function whose dimension is no smaller than τ's.

Example 1''. *If Z and W are both binary, rather than solving the system of equations implied by (11), one can instead specify a saturated model for the treatment confounding bridge function:*

$$q^*(A, Z; \tau) = \tau_0 + \tau_1 Z + \tau_2 A + \tau_3 ZA \qquad (14)$$

and estimate $\tau = (\tau_0, \tau_1, \tau_2, \tau_3)^T$ by solving (13) with $m(W, A) = (1, W, A, WA)^T$. Extension to Z and X with multiple categories is straightforward.

Example 2'. *In case of continuous (U, X, Z), result (8) suggests the model*

$$q^*(A, Z, X; \tau) = 1 + \exp\left[(-1)^A(\tau_0 + \tau_1 A + \tau_2 Z + \tau_3 X)\right]. \qquad (15)$$

If a univariate NCO W is available, we may solve (13) with $m(W, A, X) = (1, W, A, X)^T$.

2.4   Estimation and Inference

In the previous sections, we have defined the causal parameter of interest β_0 as stratum-specific log risk ratio, introduced the treatment confounding bridge function as a key ingredient to identification of β_0, and presented a strategy to estimate the treatment confounding

bridge function leveraging an available NCO. We summarize the steps of our estimation framework in Algorithm 1 and present the large-sample properties of the resulting estimator $(\widehat{\beta}, \widehat{\tau})$ in Theorem 3.

Algorithm 1 Negative control method to estimate vaccine effectiveness from a test-negative design

---

1: Identify the variables in the data according to Figure 1(e), in particular the NCEs and NCOs.
2: Estimate the treatment confounding bridge function by solving Equation (13) with a suitable parametric model $q^*(A, Z, X; \tau)$ and a user-specified function $m(W, A, X)$. Write $\widehat{\tau}$ as the resulting estimate of τ.
3: Estimate β_0 by solving Equation (9) with $\widehat{q}(A, Z, X) = q^*(A, Z, X; \widehat{\tau})$. For a user-specified nonzero one-dimensional function $c(\cdot)$, the resulting estimator of β_0 is

$$\widehat{\beta} = \log \left(\frac{\sum c(X_i) q^*(A_i, Z_i, X_i; \widehat{\tau}) A_i Y_i}{\sum c(X_i) q^*(A_i, Z_i, X_i; \widehat{\tau})(1 - A_i) Y_i} \right); \tag{16}$$

The estimated vaccine effectiveness is $\widehat{VE} = 1 - \exp(\widehat{\beta})$.

Theorem 3 (Inference based on $(\widehat{\beta}, \widehat{\tau})$)**.** *Under Assumptions 1-8 and suitable regularity conditions provided in Appendix J , the estimator $(\widehat{\beta}, \widehat{\tau})$ in Algorithm 1, or equivalently, the solution to the estimating equation $\frac{1}{n} \sum_{i=1}^{n} G_i(\beta, \tau) = 0$ is regular and asymptotically linear with the i-th influence function $IF_i(\beta, \tau) = -\Omega(\beta, \tau)^+ G_i(\beta, \tau)$, where $\Omega(\beta, \tau)^+$ denotes the Moore-Penrose inverse of $\Omega(\beta, \tau)$,*

$$G_i(\beta, \tau) = \begin{pmatrix} (-1)^{1-A_i} c(X_i) q^*(A_i, Z_i, X_i; \tau) Y_i \exp(-\beta A_i) \\ (1 - Y_i) \left[m(W_i, A_i, X_i) q^*(A_i, Z_i, X_i; \tau) - m(W_i, 1, X_i) - m(W_i, 0, X_i) \right] \end{pmatrix}$$

and $\Omega(\beta, \tau) = \left(E\left[\partial G_i(\beta, \tau) / \partial \beta^T \right], E\left[\partial G_i(\beta, \tau) / \partial \tau^T \right] \right).$

The proof follows from standard estimating equation theory (See Van der Vaart (2000) Theorem 5.21). An immediate consequence of Theorem 3 is that we may estimate the

variance-covariance matrix of $(\widehat{\beta}, \widehat{\tau})$ with

$$\widehat{\Omega}(\widehat{\beta}, \widehat{\tau})^+ \widehat{Var}(G_i(\widehat{\beta}, \widehat{\tau})) \left[\widehat{\Omega}(\widehat{\beta}, \widehat{\tau})^+ \right]^T / n, \tag{17}$$

where $\widehat{\Omega}(\beta, \tau) = \left(\hat{E} \left[\partial G_i(\beta, \tau) / \partial \beta^T \Big|_{\beta = \hat{\beta}, \tau = \hat{\tau}} \right], \hat{E} \left[\partial G_i(\beta, \tau) / \partial \tau^T \Big|_{\beta = \hat{\beta}, \tau = \hat{\tau}} \right] \right)$. Here $\hat{E}$ and $\widehat{Var}$ denote the expectation and variance with respect to the empirical distribution, respectively. A two-sided α-level Wald-type confidence interval of VE can then be obtained as $\left(1 - \exp \left(\widehat{\beta} - z_{1-\alpha/2} \sqrt{\widehat{\Sigma}_{n,1,1}} \right), 1 - \exp \left(\widehat{\beta} + z_{1-\alpha/2} \sqrt{\widehat{\Sigma}_{n,1,1}} \right) \right)$, where $\widehat{\Sigma}_{n,1,1}$ is the $(1,1)$-th entry of $\widehat{\Sigma}_n$ and $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$-th quantile of a standard normal distribution.

The estimator $\hat{\beta}$ and the above confidence interval are constructed under the assumption that the disease is rare in the target population; for non-rare diseases, $\hat{\beta}$ is in general going to be biased and the confidence interval may not be well-calibrated. However, by Corollary 1, under the null hypothesis of no vaccine effects, the estimated $q^*(A, Z, X)$ converges to the true treatment confounding bridge function and $\hat{\beta}$ is consistent for $\beta_0 = 0$. This implies that while our methods are approximately asymptotically unbiased for rare infections, they provide a valid test of no vaccine effect even if the infection is not rare.

2.5   Accounting for effect modification by measured confounders

So far we have operated under Assumption 3 that VE is constant across levels of (U, X). As we now show, this assumption can be relaxed to allow for potential effect modification with respect to X without compromising identification. This extension is particularly important because empirical evidence has indeed suggested that flu vaccine effectiveness may vary across sex and age groups (Chambers et al., 2018); and similar effect heterogeneity is of key interest in case of COVID-19 (Fernández Villalobos et al., 2021).

Instead of Assumption 3, we consider a less stringent assumption:

Assumption 9 (No effect modification by unmeasured confounders)**.**

$$P(Y = 1|A = a, U, X) = \exp(\beta_0(X)a)g(U, X) \tag{18}$$

where $\beta_0(x)$ are $g(u, x)$ are unknown functions of x and u, x respectively.

Under condition 1, Assumption 9 further implies that $\beta_0(x) = E[Y(1)|X = x]/E[Y(0)|X = x]$, i.e. the conditional causal RR as a function of x. Similar to Theorem 1, we have:

Theorem 4. *Under Assumptions 1, 2, 4, 5 and 9, for an arbitrary function $c(\cdot)$ we have that*

$$E[V_3(A, Y, Z, X; \beta_0)|S = 1] = 0,$$

where $V_3(A, Y, Z, X; \beta_0) = (-1)^{1-A}c(X)q(A, Z, X)\exp(-\beta_0(X)A)$.

The proof of Theorem 4 is identical to that of Theorem 1 with $\beta_0 A$ replaced with $\beta_0(X)A$. Identification and estimation of the treatment confounding bridge function are also essentially identical to that of Corollary 2. Therefore, it is straightforward to extend Algorithm 1 to allow effect modification by measured confounders. We describe the algorithm and the large sample properties of the resulting estimator in Appendix K.

2.6 Estimating VE under treatment-induced selection into TND sample

Thus far, unbiasedness of the estimating function V_0 has crucially relied on Assumption 2 that A does not have a direct effect on S. In some settings, the assumption may be violated if an infected person who is vaccinated is on average more likely to present to the ER than an unvaccinated infected person with similar symptoms, so that treatment or

vaccination-induced selection into the TND sample is said to be present. In such settings, both estimators $\hat{\beta}$ and $\hat{\beta}(X)$ produced by Algorithms 1 and 2 may be severely biased because Assumption 2 may no longer be valid. Crucially, we note that this form of selection bias can be present even in context of a randomized trial in which vaccination/treatment is assigned completely at random, if the outcome is ascertained using a TND, for example in the cluster-randomized test-negative design studies of community-level dengue intervention effectiveness Anders et al. (2018), Dufault and Jewell (2020), Jewell et al. (2019), and Wang et al. (2022). In this Section, we provide sufficient conditions for identification under treatment-induced selection. In this vein, consider the following assumptions:

Assumption 2′. $P(S = 1|A = a, Y = 1, U, X)/P(S = 1|A = a, Y = 0, U, X) = \exp(h(U, X))$ *for* $a = 0, 1$.

That is, the risk ratio association between infection status and selection into the TND sample is independent of vaccination status. Furthermore,

Assumption 3′. *(No effect modification by confounders on the OR scale).*

$$\frac{P(Y = 1|A = 1, U, X)/P(Y = 0|A = 1, U, X)}{P(Y = 1|A = 0, U, X)/P(Y = 0|A = 0, U, X)} = \exp(\beta_0').$$

Recall that Assumption 3 posited a constant vaccination causal effect on the RR scale across levels of (U, X), while Assumption 3′ posits that the corresponding causal effect on the odds ratio scale is constant across levels of (U, X). In case of a rare infection in the target population, the OR and RR are approximately equal, in which case VE is well approximated by $1 - OR$.

Furthermore, identification relies on the following modified definition of a treatment confounding bridge function:

Assumption 5′. *There exists a treatment confounding bridge function $\tilde{q}$ such that for $a = 0, 1$,*

$$E[\tilde{q}(a, Z, X)|A = a, U, X] = 1/P(A = a|U, X, Y = 0, S = 1) \qquad almost\ surely. \qquad (19)$$

Note that if the infection is rare in the target population in the sense of Assumption 7, then the treatment confounding bridge function defined in Assumption 5 in Section 2.3.1 satisfies (19) approximately .

We now introduce the identification of the OR with the following theorem:

Theorem 1′. *Under Assumptions 1, 2′, 3′, 4 and 5′, we have*

$$E[\tilde{V}_1(A, Y, Z, X; \beta_0') \mid U, X, S = 1] = 0$$

where $\tilde{V}_1(A, Y, Z, X; \beta)$ is the same as V_1 defined in Theorem 1 except with $\tilde{q}$ replacing q.

Importantly, the theorem establishes that the estimating function V_1 previously developed in the paper can under certain conditions, remain unbiased for the odds ratio association of vaccination with testing positive for the infection, even in the presence of treatment-induced selection into the TND sample. We leave the proof of Theorem 1′ to Appendix L.

Estimation of the treatment confounding bridge function $\tilde{q}(A, Z, X)$ requires a negative control outcome that satisfies:

Assumption 6′. *(NCO Independence Conditions)* $W \perp\!\!\!\perp (A, Z, S)|U, X, Y$.

In addition to Assumptions 6, this last assumption requires that neither Y nor S is a causal effect of W. Figure 1(f) illustrates a DAG that satisfies our assumptions regarding

(Z, W). As can be verified in the graph, Assumption 6′ is needed to ensure that collider stratification bias induced by the path $A \to [S = 1] \leftarrow W$ upon conditioning on $S = 1$ is no longer present. Identification of the function $\tilde{q}$ is given below:

Theorem 2′. *Under assumptions 4, 5′ and 6′, for $a = 0, 1$ we have that*

$$E[\tilde{q}(a, Z, X)|A = a, W, X, Y = 0, S = 1] = 1/P(A = a|W, X, Y = 0, S = 1)$$

We prove Theorem 2′ in Appendix M. As a result of Theorem 2′, the parameters in the treatment confounding bridge function can be estimated by solving moment equation (13).

In summary, the above discussion suggests that one can continue to use Algorithm 1 to estimate VE in presence of treatment induced selection bias, albeit on the OR scale and under a modified set of negative control conditions. Algorithm 2 can similarly be justified under treatment-induced selection with assumptions in this section, except that β'_0 in Assumption 3′ is replaced by the conditional log RR $\beta'_0(X)$.

As a side note, Assumption 2′ automatically holds under Assumption 2, and hence the above results in this section also apply to the setting in previous sections that is illustrated in Figure 1(e). We present this statement in the following corollary.

Corollary 3. *Under Assumptions 1, 2, 3′, 4 and 5′, we have*

$$E[\tilde{V}_1(A, Y, Z, X; \beta'_0) \mid U, XS = 1] = 0.$$

With Assumption 2, the treatment confounding bridge function $\tilde{q}$ can be estimated by solving the moment equation (13) either under under Assumption 6′ and 8, or Assumptions 6, 7 and 8 as an approximation under the rare disease assumption. Corollary 3 leads to an interesting observation: under the treatment-independent sampling (Assumption 2),

the estimator $\hat{\beta}$ from Algorithm 1 can be viewed as either log RR or log OR, depending on the set of assumptions made.

3 Simulation Study

To assess the empirical performance of our proposed method, we consider two settings with different types of confounding and negative control variables, and perform corresponding simulation studies.

In the first setting, we consider no measured confounder, a binary unmeasured confounder U, a binary NCE Z and a binary NCO W. To trigger selection among subjects with $Y = 0$, we let D be a binary indicator of the presence of other flu like illnesses. The treatment confounding bridge function is thus given by (7). We assume the distribution of Y is Bernoulli with a log-linear risk model: $Y|A, U \sim Bernoulli(\exp(\eta_{0Y} + \beta_0 A + \eta_{UY} U))$. We consider values of β_0 to be -1.609, -0.693, -0.357 or 0, corresponding to a risk ratio of 0.2, 0.5, 0.7 or 1. We assume the selection S only equals one with nonzero probability if at least one of Y, W and D equals one, and is independent of A and Z conditional on other variables. The resulting prevalence of Y in the target population is 0.75% among the unvaccinated individuals and 0.55%, 0.65%, 0.72% or 0.75% among the vaccinated individuals, corresponding to four values of β_0. Next, we consider a setting where X, U, Z and W are all univariate continuous variables. We generate the infection outcome using a log-linear model

$$Y|A, U, X \sim Bernoulli(\exp(\mu_{0Y} + \beta_0 A + \mu_{UY} U + \mu_{XY} X + \mu_{UXY} UX)).$$

We generate A and Z following Example 2 in Section 2.3. As such the treatment con-

founding bridge function is given by Equation (8). The probability of $S = 1$ is 1 only if at least one of Y and D is nonzero. The resulting prevalence of Y in the target population is 0.34% among the unvaccinated individuals and 0.24%, 0.28%, 0.31% or 0.34% among the vaccinated individuals, corresponding to four values of β_0. Appendix N and O give more details on the data-generating mechanism for the two settings.

In each scenario, we simulate a target population of size $N = 7,000,000$ and implement $1,000$ simulation iterations. For both settings, we evaluate the performance of three estimators for β_0:

- NC estimator: our proposed estimator given by Algorithm 1. In the first setting, we use a saturated parametric model (14) for the treatment confounding bridge function in the first setting, with $m(W, A) = (1, W, A, WA)^T$; in the second setting, we use model (15) and $m(W, A, X) = (1, W, A, X)^T$. We set $c(X) = 1$ in both settings.

- NC-Oracle estimator: the estimated treatment confounding bridge function in Algorithm 1 is only an approximation under Assumption 7, whose bias may affect the estimation for β_0, as derived in Appendix I. We therefore include NC-Oracle estimator that uses the true treatment confounding bridge function $q(A, Z, X)$. Appendices E include derivation of the true treatment confounding bridge function under the continuous (X, U, Z, W) setting.

- Logistic regression: we also consider a logistic regression model of Y on A (and X in the second setting), overlooking the unmeasured confounders U. This is a common choice for covariate adjusted analyses of test-negative designs but ignores biases caused by U (Bond, Sullivan, and Cowling, 2016). We comment in Appendix P that the estimator is appropriate in the absence of unmeasured confounders except for potential model misspecification.

We note that the target parameter β_0 for NC estimator and NC-Oracle estimator is log causal RR, while logistic regression gives log causal OR. However, the two parameters are approximately equal under Assumption 7 where the infection risk is low in the target population.

Figure 2 shows the bias of three estimators considered and the coverage of their 95% confidence intervals. In both settings, both NC and NC-Oracle are essentially unbiased whereas logistic regression gives a biased estimate in all scenarios. NC-Oracle exhibits slightly higher precision than NC, which implies that estimating the treatment confounding bridge function in the TND is only slightly more variable. The 95% confidence intervals for NC and NC-Oracle both achieve nominal coverage, whereas logistic regression-based confidence intervals under-cover severely. We repeated the simulation under a non-rare disease setting in Appendix Q. In such scenarios, while NC-Oracle estimator is still unbiased with calibrated 95% confidence intervals, the NC estimator is biased in general except when $\beta_0 = 0$. We conclude that the proposed NC estimator is unbiased of the log causal RR either under a rare disease setting or under a non-rare disease setting with no vaccine effect.

4 Application

We applied our proposed method to a TND study of COVID-19 VE of two-dose Moderna vaccine (mRNA-1273), two-dose Pfizer-BioNTech vaccine (BNT162b2), and single-dose Johnson & Johnson's Janssen vaccine (Ad26.COV2.S) against COVID-19 infection nested in the University of Michigan Health System. The selected study sample includes patients who interacted with the University of Michigan Health System and experienced COVID-19 symptoms, had suspected exposure to COVID-19 virus, or sought to screen for COVID-19 infection, between April 5, 2021 and December 7, 2021. In addition, the selected
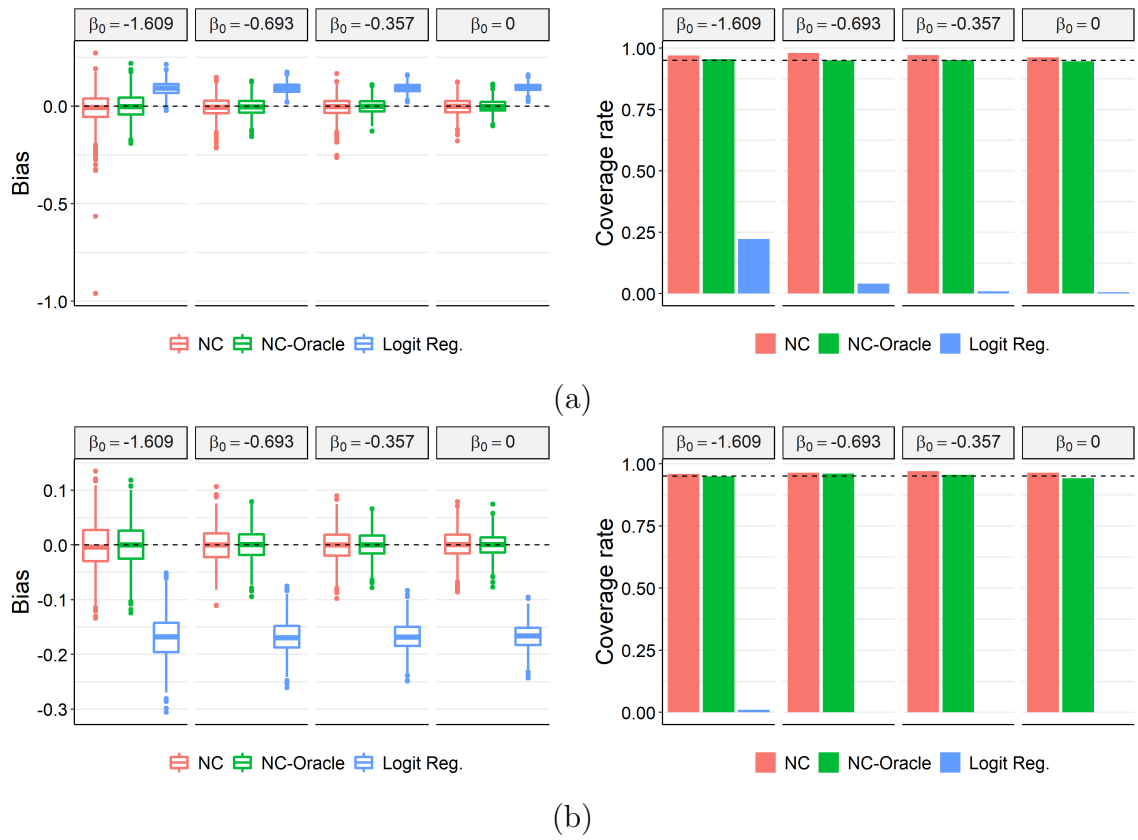
Figure 2: Bias (left) and coverage rates of 95% confidence interval (right) for the oracle estimator (NC-Oracle), GMM estimator (NC-GMM) and logistic regression (Logit Reg.) with a (a) binary or (b) continuous unmeasured confounder.

test-positive subjects had at least one positive lab tests for COVID-19 infection after April 5. Vaccination history was obtained through electronic health records. A study subject was considered fully vaccinated if they received at least one dose of Johnson & Johnson's Janssen vaccine or at least two doses of Moderna or Pfizer-BioNTech vaccine. If a subject tested positive before or within 14 days after their first dose of Janssen vaccine or within 14 days after their second dose of Moderna or Pfizer-BioNTech vaccine, they were considered unvaccinated (Moline et al., 2021).

We took immunization visits before December 2020 as NCE since COVID-19 vaccines were not available before December 2020 and immunization before was unlikely to affect the risk of COVID-19 infection; nor that of the selected NCOs we describe next. For

NCO, we selected a binary indicator of having at least one of the following "negative control outcome" conditions after April 5, 2021: arm/leg cellulitis, eye/ear disorder, gastro-esophageal disease, atopic dermatitis, and injuries. Such candidate NCE and NCO are likely to satisfy the requisite conditional independence conditions for them to be valid negative control variables and to be related to a patient's latent HSB. We adjusted for age groups (<18, between 18 and 60, or ≥ 60), gender, race (white or non-white), Charlson comorbidity score ≥ 3, and the calendar month of a test-positive subject's first positive COVID test or a test-negative subject's last COVID test. Table S3 in Appendix R summarizes the distribution of negative control variables, demographic variables and COVID-19 infection among vaccinated and unvaccinated subjects.

Because NCE is expected not to be associated with either the outcome or NCE in a fully adjusted analysis unless there is unmeasured confounding, we first fit regression models to detect presence of residual confounding bias. Conditioning on the baseline covariates, in both vaccinated and unvaccinated groups, NCE is significantly associated with COVID-19 infection ($p < 0.001$) and NCO ($p < 0.001$) in corresponding adjusted logistic regression models, suggesting the presence of hidden biases (See Appendix R Table S4, S5).

We implemented Algorithm 1 to estimate VE. We specified a linear model for the treatment confounding bridge function with an interaction term between COVID-19 vaccination and the NCE, and set the function m to include one (for an intercept term), COVID-19 vaccination, the NCO, and baseline covariates, as well as all two-way interactions. For comparison, we also implemented a logistic regression model, which gives an unbiased estimate of causal odds ratio under the no unmeasured confounding assumption, adjusting for gender and age groups. In this case, The VE can be approximated by one minus the odds ratio of COVID-19 infection against vaccination, as COVID-19 infection rate is known to

be low across strata in the target population.

The double negative control Algorithm 1 estimated the VE against lab-confirmed COVID-19 infection were 71.6% (95% CI: 68.9%, 74.0%) for the two-dose Pfizer-BioNTech vaccine, 86.8% (95% CI: 84.7%, 88.6%) for the two-dose Moderna vaccine, and 66.0% (95% CI: 54.5%, 74.6%) for the one-dose Johnson & Johnson's Janssen vaccine. The logistic regression estimates for the same VE were 67.0% (95% CI: 64.5%, 69.4%) for the Pfizer-BioNTech vaccine, 75.1% (95% CI, 72.5%, 77.5%) for the Moderna vaccine, and 56.0% (95%: 47.9%, 62.9%) for the Janssen vaccine. There is significant evidence of hidden bias as summarized in Tables S4 and S5. The VE estimates with our double NC method are notably higher than those given by the standard logistic regression for all three regimens.

Recent observational studies estimated real-world mRNA COVID-19 vaccine (Pfizer-BioNTech and Moderna) effectiveness ranging between 80% and 98% against lab confirmed SARS-COV-2 infection of different variants (**bruxvoort2021effectiveness**; **israel2021elapsed**; Dagan et al., 2021). Our NC approach provided VE estimates that are close to the previous studies. We hypothesize that the standard logistic regression underestimates the VE by overlooking the confounding bias due to HSB and related factors, which our proposed double NC approach appears to reduce to some extent.

5 Discussion

In this article, we have introduced a statistical method for estimating vaccine effectiveness in a test-negative design. The approach leverages negative control variables to account for hidden bias due to residual confounding and/or selection mechanism into the TND sample. Negative control variables abound in practice, such as vaccination history which is routinely collected in insurance claims and electronic health records. Hence the proposed

method may be particularly useful in such real world settings to obtain improved estimates of vaccine effectiveness.

The TND is a challenging setting in causal inference where selection bias and unmeasured confounding co-exist, selection is outcome-dependent, and unmeasured confounders also impact selection. As a result, the causal effect of interest is in general not identified from such studies (Cai and Kuroki, 2012). Nevertheless, we establish that progress can be made under a semiparametric multiplicative model, provided the outcome is rare in the target population, and double negative control variables are available. To this end, this article showcases the potential power of negative control methods and proximal causal inference in epidemiologic research (Shi, Miao, and Tchetgen Tchetgen, 2020; Tchetgen Tchetgen et al., 2020).

We focused on the outpatient TND, where recruitment is restricted to subjects who seek care voluntarily. TNDs have also been applied to inpatient settings for studying VE against, for example, flu hospitalization (Feng, Cowling, and Sullivan, 2016; Foppa et al., 2016). In inpatient TNDs, differential access to healthcare and underlying health characteristics between vaccinated and unvaccinated subjects are likely the main causes of confounding bias (Feng, Cowling, and Sullivan, 2016). Our methods are still applicable in such settings, but negative control variables should be selected to be relevant to the source of unmeasured confounding mechanism. For example, previous vaccination and hospitalization outside the flu season or hospitalization due to other flu-like illnesses are viable candidate NCE and NCO, respectively (Jackson et al., 2006).

Our approach is suitable for post-market TND studies where real-world vaccine effectiveness is of interest and vaccination history is obtained retrospectively, possibly through electronic health records. For vaccine efficacy in a controlled trial setting, Wang et al. (2022)

recently developed estimation and inference of RR in cluster-randomized TND, aiming to correct for bias due to differential HSB induced by the intervention being unblinded. Because of randomization, they considered HSB as a post-treatment variable and proposed a log-contrast estimator which corrects for selection bias by leveraging a valid test-negative outcome, under an assumption that either (i) the vaccine does not have a causal effect in the population, and the causal impact of vaccination on selection is equal for test-positive and -negative subsamples; or (ii) among care seekers, the incidence of test-negative outcomes does not differ between vaccinated and unvaccinated, and the intervention effect among care seekers is generalizable to the whole population. We note that even under randomization, identification conditions given in Section 2.6 are neither stronger nor weaker than those of Wang et al. (2022) described above, as neither set of assumptions appear to imply the other. An important advantage of our proposed methods is that they can be used to account for selection bias in a TND study irrespective of randomization.

Our methods target RR as a measure of VE instead of the more common OR (Jackson et al., 2006; Sullivan, Tchetgen Tchetgen, and Cowling, 2016). These two measures are approximately equal for rare infections. Schnitzer (2022) recently considered estimation of a marginal causal RR in the TND sample and justified the use of an inverse probability of treatment weighted (IPTW) estimator in a setting in which an unmeasured common cause of infection and selection into the TND sample does not cause vaccination (and thus there is no unmeasured confounding). Instead, our methods allow for an unmeasured common cause of vaccination, infection and selection into the TND sample; however in order to estimate a causal RR, we invoke both, an assumptions of no effect modification by an unmeasured confounder, and a rare-disease condition. As we establish, the latter assumption is not needed if there is no vaccine effect against infection outcome. In Section 2.6, we es-

tablish that under a homogeneous OR vaccine effect measure condition, and an alternative definition of the treatment bridge function, our methods can identify a causal effect of the vaccine on the odds ratio scale without invoking the rare disease condition.

Throughout the article, we have assumed diagnostic tests are accurate and individuals who seek care are sparsely distributed, such that the vaccination of a given subject in the TND sample does not protect another study subject from infection, , i.e. there is no interference in the TND sample, a common assumption in TND literature. This assumption may be violated if members of the same households present in the ER in which case block interference must be accounted using results from interference literature (Hudgens and Halloran, 2008; Tchetgen Tchetgen and VanderWeele, 2012). Sensitivity analysis may be considered to evaluate how violation of these assumptions can potentially bias inferences about VE.

Acknowledgements

A Proof of Proposition 1

The result immediately follows if we can show that

$$E\left[(-1)^{1-A}\frac{1}{P(A|U,X)}\exp(-\beta_0 A)\bigg|U,X,Y=1,S=1\right]=0. \tag{20}$$

By Assumption 2, the left-hand side of (20) equals

$$E\left[(-1)^{1-A}\frac{1}{P(A|U,X)}\exp(-\beta_0 A)\bigg|U,X,Y=1\right].$$

We further have

$$E\left[(-1)^{1-A}\frac{1}{P(A|U,X)}\exp(-\beta_0 A)\bigg|U,X,Y=1\right]$$

$$=\sum_{a=0}^{1}(-1)^{1-a}\frac{1}{P(A=a|U,X)}\exp(-\beta_0 a)P(A=a|U,X,Y=1)$$

$$=\sum_{a=0}^{1}(-1)^{1-a}\frac{1}{P(A=a|U,X)}\exp(-\beta_0 a)\frac{P(A=a|U,X)P(Y=1|A=a,U,X)}{P(Y=1|U,X)}$$

$$\overset{A.3}{=}\sum_{a=0}^{1}(-1)^{1-a}\exp(-\beta_0 a)\frac{\exp(\beta_0 a)P(Y=1|A=0,U,X)}{P(Y=1|U,X)}$$

$$=\sum_{a=0}^{1}(-1)^{1-a}\frac{P(Y=1|A=0,U,X)}{P(Y=1|U,X)}$$

$$=0.$$

40

B   Existence of solutions to (6)

In this section, we provide the conditions of existence of solutions to (6). The conditions for (11) can be similarly derived. The results in this section directly adapted from Appendix B of Cui et al. (2020).

Let $L^2\{F(t)\}$ denote the Hilbert space of all square-integrable functions of t with respect to distribution function $F(t)$, equiped with inner product $\langle g_1, g_2 \rangle = \int g_1(t)g_2(t)dF(t)$. Let $T_{a,x}$ denote the operator $L^2\{F(z|a, x)\} \to L^2\{F(u|a, x)\}$, $T_{a,x}q = E[q(Z)|A = a, U = u, X = x]$ and let $(\lambda_{a,x,n}, \varphi_{a,x,n}, \phi_{a,x,n})$ denote a singular value decomposition of $T_{a,x}$. The solution to (6) exists if:

(1) $\int \int f(z|a, u, x)f(u|a, z, x)dzdu < \infty$;

(2) $\int P^{-2}(A = a|U = u, X = x)f(u|a, x)du < \infty$;

(3) $\sum_{n=1}^{\infty} \lambda_{a,x,n}^{-2}|\langle P^{-1}(A = a|U = u, X = x), \phi_{a,x,n}\rangle|^2 < \infty$.

C   Proof of Theorem 1

It suffices to show that

$$E[(-1)^{1-A}q(A, Z, X)\exp(-\beta_0 A)|U, X, Y = 1, S = 1] = 0.$$

By Assumption 2, the left-hand side is

$$E[e^{-\beta_0 A}(-1)^{1-A}q(A, Z, X)|U, X, Y = 1]$$

$$= E\{e^{-\beta_0 A}(-1)^{1-A}E[q(A, Z, X)|A, U, X, Y = 1]|U, X, Y = 1\}$$

$$\overset{A.4}{=} E\{e^{-\beta_0 A}(-1)^{1-A}E[q(A, Z, X)|A, U, X]|U, X, Y = 1\}$$

$$\overset{A.5}{=} E\{e^{-\beta_0 A}(-1)^{1-A}\frac{1}{P(A|U, X)}|U, X, Y = 1\}$$

$$= 0$$

The last equality is proved in Appendix A.

D Treatmeng bridge function and estimation with categorical NCE, NCO and unmeasured confounders.

We consider a categorical unmeasured confounder U with categories $u_1, \ldots, u_J$ and NCE Z with categories $z_1, \ldots, z_I$. We assume there are no other covariates X. We write $p_{ia.k} = P(Z = z_i, A = a | U = u_j)$ for $i = 1, \ldots, I$, $a = 0, 1$, and $j = 1, \ldots, J$.

Similar to before, the treatment confounding bridge function $q(a, z)$ should satisfy

$$\sum_{i=1}^{I} p_{ia.j} q(a, z_i) = 1 \tag{21}$$

for all i, j. Therefore, any solution of the following equation system, if exists, is a treatment confounding bridge function:

$$p_{1a.1} q(a, z_1) + p_{2a.1} q(a, z_2) + \cdots + p_{Ia.1} q(a, z_I) = 1$$
$$p_{1a.2} q(a, z_1) + p_{2a.2} q(a, z_2) + \cdots + p_{Ia.2} q(a, z_I) = 1$$
$$\cdots \tag{22}$$
$$p_{1a.J} q(a, z_1) + p_{2a.J} q(a, z_2) + \cdots + p_{Ia.J} q(a, z_I) = 1$$

for $a = 0, 1$. We denote the probability matrix

$$P_{Z,A|U} = \begin{pmatrix} p_{1a.1} & p_{2a.1} & \cdots & p_{Ia.1} \\ p_{1a.2} & p_{2a.2} & \cdots & p_{Ia.2} \\ \vdots & & & \vdots \\ p_{1a.J} & p_{2a.J} & \cdots & p_{Ia.J} \end{pmatrix}. \tag{23}$$

A treatment confounding bridge function exists if the matrix $P_{Z,A|U}$ is invertible.

Suppose besides a categorical NCE Z with levels $z_1, \ldots, z_I$, we also have a categorical

NCO W with levels $w_1, \ldots, w_K$. The integral equation (11) is equivalent to the linear equation system

$$p'_{1a.1}q^*(a, z_1) + p'_{2a.1}q^*(a, z_2) + \cdots + p'_{Ia.1}q^*(a, z_I) = 1$$

$$p'_{1a.2}q^*(a, z_1) + p'_{2a.2}q^*(a, z_2) + \cdots + p'_{Ia.2}q^*(a, z_I) = 1$$

$$\cdots$$

$$p'_{1a.K}q^*(a, z_1) + p'_{2a.K}q^*(a, z_2) + \cdots + p'_{Ia.K}q^*(a, z_I) = 1$$

(24)

for $a = 0, 1$, where $p'_{ia.k} = P(Z = z_i, A = a | W = w_k, Y = 0, S = 1)$. We denote that matrix

$$P'(a) = \begin{pmatrix} p'_{1a.1} & p'_{2a.1} & \cdots & p'_{Ia.1} \\ p'_{1a.2} & p'_{2a.2} & \cdots & p'_{Ia.2} \\ \vdots & & & \vdots \\ p'_{1a.K} & p'_{2a.K} & \cdots & p'_{Ia.K} \end{pmatrix}.$$

(25)

Then Assumption 8 is equivalent to the condition that $P'(a)$ is invertible with $I = K$ for $a = 0, 1$, in which case (11) has a unique solution

$$(q^*(a, z_1), \ldots, q^*(a, z_I))^T = [P'(a)]^{-1}\mathbf{1}_I.$$

The probabilities $p'_{ia.k}$'s can all be estimated from the study data.

E   Derivation of the treatment confounding bridge function in Example 2

By Assumption 5, the treatment confounding bridge function $q(A, Z, X)$ should satisfy

$$E[q(A, Z, X)|U = u, A = a, X = x] = \frac{1}{P(A = a|U = u, X = x)}$$

for all a, u and X.

We write $q(A, Z, X) = 1 + r(Z, A, X)$, then

$$E[r(Z, a, X)|U, A = a, X] = \frac{1 - P(A = a|U, X)}{P(A = a|U, X)}.$$

Consider

$$r(Z, A, X) = \exp((-1)^A(\tau_0 + \tau_1 A + \tau_2 Z + \tau_3 X)),$$

then

$$
\begin{aligned}
&\frac{P(A = 1|U = u, X = x)}{1 - P(A = 1|U = u, X = x)} \\
&= \int r(z, 0, x) f(z|u, x, A = 0) dz \\
&= \int \exp(\tau_0 + \tau_2 z + \tau_3 x) \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(-\frac{(z - \mu_{0z} - \mu_{UZ}u - \mu_{XZ}x)^2}{2\sigma_z^2}\right) dz \\
&= \exp(\tau_0 + \tau_3 x) \exp\left(\tau_2(\mu_{0Z} + \mu_{UZ}u + \mu_{XZ}x) + \frac{\sigma_z^2\tau_2^2}{2}\right) \\
&= \exp\left(\tau_0 + \tau_2\mu_{0Z} + \frac{\sigma_z^2\tau_2^2}{2} + \tau_2\mu_{UZ}u + (\tau_3 + \mu_{XZ}\tau_2)x\right)
\end{aligned}
$$

and

$$\frac{1 - P(A = 1|U = u, X = x)}{P(A = 1|U = u, X = x)}$$

$$= \int r(z, 1, x) f(z|u, x, A = 1) dz$$

$$= \int \exp(-\tau_0 - \tau_1 - \tau_2 z - \tau_3 x) \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(-\frac{(z - \mu_{0Z} - \mu_{AZ} - \mu_{UZ}u - \mu_{XZ}x)^2}{2\sigma_z^2}\right) dz$$

$$= \exp(-\tau_0 - \tau_1 - \tau_3 x) \exp\left(-\tau_2(\mu_{0Z} + \mu_{AZ} + \mu_{UZ}u + \mu_{XZ}x) + \frac{\sigma_z^2\tau_2^2}{2}\right)$$

$$= \exp\left(-\tau_0 - \tau_1 - \tau_2\mu_{0Z} - \tau_2\mu_{AZ} + \frac{\sigma_z^2\tau_2^2}{2} - \tau_2\mu_{UZ}u - (\tau_3 + \tau_2\mu_{XZ})x\right)$$

This requires that

$$\tau_1 + \tau_2\mu_{AZ} = \sigma_z^2\tau_2^2$$

Because

$$P(A = 1|U, X) = \mathrm{expit}(\mu_{0A} + \mu_{UA}U + \mu_{XA}X),$$

we conclude the parameters in the bridge function are

$$\tau_2 = \mu_{UA}/\mu_{UZ},$$

$$\tau_3 = \mu_{XA} - \mu_{XZ}\tau_2 = \mu_{XA} - \mu_{XZ}\mu_{UA}/\mu_{UZ}$$

$$\tau_1 = \sigma_z^2\tau_2^2 - \tau_2\mu_{AZ} = \frac{\sigma_z^2\mu_{UA}^2}{\mu_{UZ}^2} - \frac{\mu_{UA}\mu_{AZ}}{\mu_{UZ}}$$

$$\tau_0 = \mu_{0A} - \tau_2\mu_{0Z} - \frac{\sigma_z^2\alpha_2^2}{2} = \mu_{0A} - \frac{\mu_{UA}\mu_{0Z}}{\mu_{UZ}} - \frac{\sigma_z^2\mu_{UA}^2}{2\mu_{UZ}^2}$$

F   Proof of Theorem 2 and Corollary 1

We first introduce a few properties due to the rare disease assumption:

Lemma 1. *Under Assumptions 2, 4, 6 and 7, for every a, w, u and x, we have*

(a)

$$P(Y = 1|W = w, U = u, X = x) < \delta, \qquad P(Y = 1|A = a, U = u, X = x) < \delta,$$

$$P(Y = 1|U = u, X = x) < \delta.$$

(b)

$$1 - \delta < \frac{P(A = a|U = u, X = x, Y = 0, S = 1)}{P(A = a|U = u, X = x)} < \frac{1}{1 - \delta}$$

(c)

$$(1 - \delta)^2 < \frac{P(A = a|W = w, U = u, X = x, Y = 0, S = 1)}{P(A = a|U = u, X = x, Y = 0, S = 1)} < \frac{1}{(1 - \delta)^2}.$$

(d)

$$f(z|A = a, U = u, X = x) = f(z|W = w, A = a, U = u, X = x, Y = 0, S = 1).$$

Proof.   (a) For every a, w, u and x, we have

$$P(Y = 1|W = w, U = u, X = x)$$

$$= \sum_a P(Y = 1|A = a, W = w, U = u, X = x)P(A = a|W = w, U = u, X = x)$$

$$\overset{A.7}{<} \delta \sum_a P(A = a|W = w, U = u, X = x)$$

$$= \delta.$$

The rest follows similarly.

(b) For every a, u and x, we have

$$P(A = a|U = u, X = x, Y = 0) = P(A = a|U = u, X = x) \times \frac{P(Y = 0|A = a, U = u, X = x)}{P(Y = 0|U = u, X = x)}.$$

By Lemma 1(a), we have

$$1 - \delta < \frac{P(Y = 0|A = a, U = u, X = x)}{P(Y = 0|U = u, X = x)} = \frac{P(A = a|U = u, X = x, Y = 0)}{P(A = a|U = u, X = x)} < \frac{1}{1 - \delta}.$$

The result follows by noticing that $P(A = a|U = u, X = x, Y = 0) = P(A = a|U = u, X = x, Y = 0, S = 1)$ due to Assumption 2.

(c) For every a, w, u and x, we have

$$P(A = a|W = w, U = u, X = x, Y = 0)$$

$$= P(A = a|W = w, U = u, X = x) \times \frac{P(Y = 0|A = a, W = w, U = u, X = x)}{P(Y = 0|W = w, U = u, X = x)}$$

$$\overset{A.6(a)}{=} P(A = a|U = u, X = x) \times \frac{P(Y = 0|A = a, W = w, U = u, X = x)}{P(Y = 0|W = w, U = u, X = x)}$$

$$= P(A = a|U = u, X = x, Y = 0) \times \frac{P(Y = 0|U = u, X = x)}{P(Y = 0|A = a, U = u, X = x)}$$

$$\times \frac{P(Y = 0|A = a, W = w, U = u, X = x)}{P(Y = 0|W = w, U = u, X = x)}$$

By Lemma 1(a), we have

$$1 - \delta < \frac{P(Y = 0|U = u, X = x)}{P(Y = 0|A = a, U = u, X = x)} < \frac{1}{1 - \delta}$$

and

$$1 - \delta < \frac{P(Y = 0|A = a, W = w, U = u, X = x)}{P(Y = 0|W = w, U = u, X = x)} < \frac{1}{1 - \delta}.$$

We therefore have

$$(1 - \delta)^2 < \frac{P(A = a|W = w, U = u, X = x, Y = 0)}{P(A = a|U = u, X = x, Y = 0)} < \frac{1}{(1 - \delta)^2}.$$

Finally, by Assumptions 2 and 6(c), we have

$$P(A = a|U = u, X = x, Y = 0) = P(A = a|U = u, X = x, Y = 0, S = 1),$$

$$P(A = a|W = w, U = u, X = x, Y = 0) = P(A = a|W = w, U = u, X = x, Y = 0, S = 1).$$

We conclude that

$$(1 - \delta)^2 < \frac{P(A = a|W = w, U = u, X = x, Y = 0, S = 1)}{P(A = a|U = u, X = x, Y = 0, S = 1)} < \frac{1}{(1 - \delta)^2}.$$

(d)

$$f(z|W = w, A = a, U = u, X = x, Y = 0, S = 1)$$

$$\overset{A.6(c)}{=} f(z|W = w, A = a, U = u, X = x, Y = 0)$$

$$\overset{A.6(b)}{=} f(z|A = a, U = u, X = x)$$

$\square$

Therefore, we have

$$E\{q(a, Z, X) | A = a, W, X, Y = 0, S = 1\}$$

$$= \int q(a, z, X) f(z | A = a, W, X, Y = 0, S = 1) \, \mathrm{d}z$$

$$= \int \int q(a, z, X) f(z | A = a, W, U = u, X, Y = 0, S = 1) f(u | A = a, W, X, Y = 0, S = 1) \, \mathrm{d}z \, \mathrm{d}u$$

$$\overset{L.1(d)}{=} \int \left\{ \int q(a, z, X) f(z | A = a, U = u, X) \, \mathrm{d}u \right\} f(u | A = a, W, X, Y = 0, S = 1) \, \mathrm{d}z$$

$$\overset{A.5}{=} \int \frac{1}{P(A = a | U = u, X)} f(u | A = a, W, X, Y = 0, S = 1) du$$

$$\overset{L.1(b)}{<} \frac{1}{1 - \delta} \int \frac{1}{P(A = a | U = u, X, Y = 0, S = 1)} f(u | A = a, W, X, Y = 0, S = 1) du$$

$$\overset{L.1(c)}{<} \frac{1}{(1 - \delta)^3} \int \frac{1}{P(A = a | W, U = u, X, Y = 0, S = 1)} f(u | A = a, W, X, Y = 0, S = 1) du$$

$$= \frac{1}{(1 - \delta)^3} \int \frac{f(u | W, X, Y = 0, S = 1) f(u | A = a, W, X, Y = 0, S = 1)}{P(A = a | W, X, Y = 0, S = 1) f(u | A = a, W, X, Y = 0, S = 1)} \, du$$

$$= \frac{1}{(1 - \delta)^3} \frac{1}{P(A = a | W, X, Y = 0, S = 1)} \int f(u | W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$= \frac{1}{(1 - \delta)^3} \frac{1}{P(A = a | W, X, Y = 0, S = 1)}$$

and

$$E\{q(a, Z, X)|A = a, W, X, Y = 0, S = 1\}$$

$$= \int \frac{1}{P(A = a|U = u, X)} f(u|A = a, W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$\overset{L.1(b)}{>} (1 - \delta) \int \frac{1}{P(A = a|U = u, X, Y = 0, S = 1)} f(u|A = a, W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$\overset{L.1(c)}{>} (1 - \delta)^3 \int \frac{1}{P(A = a|W, U = u, X, Y = 0, S = 1)} f(u|A = a, W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$= (1 - \delta)^3 \int \frac{f(u|W, X, Y = 0, S = 1) f(u|A = a, W, X, Y = 0, S = 1)}{P(A = a|W, X, Y = 0, S = 1) f(u|A = a, W, X, Y = 0, S = 1)} \, \mathrm{d}u$$

$$= (1 - \delta)^3 \frac{1}{P(A = a|W, X, Y = 0, S = 1)} \int f(u|W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$= (1 - \delta)^3 \frac{1}{P(A = a|W, X, Y = 0, S = 1)}.$$

To prove Corollary 1, we have

$$E\{q(a, Z, X)|A = a, W, X, Y = 0, S = 1\}$$

$$= \int \frac{1}{P(A = a|U = u, X)} f(u|A = a, W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$= \int \frac{P(A = a|U = u, X, Y = 0, W, S = 1)}{P(A = a|U = u, X) P(A = a|W, X, Y = 0, S = 1)} f(u|W, X, Y = 0, S = 1) \, \mathrm{d}u$$

$$\overset{A.6}{=} \int \frac{P(A = a|U = u, X, Y = 0, W)}{P(A = a|U = u, X) P(A = a|W, X, Y = 0, S = 1)} f(u|W, X, Y = 0, S = 1) \, \mathrm{d}u$$

If $A \perp\!\!\!\perp Y|U, X$, then together with Assumption 6 we have $P(A = a|U = u, X, Y = 0, W) = P(A = a|U = u, X)$, whereby the above equals

$$\frac{1}{P(A = a|W, X, Y = 0, S = 1)} \int f(u|W, X, Y = 0, S = 1) du = \frac{1}{P(A = a|W, X, Y = 0, S = 1)}.$$

G Discussion of Assumption 8

Similar to (6), Equation (11) defines a Fredholm integral equation of the first kind, yet only involves observed data. Although a treatment confounding bridge function $q(A, Z, X)$ must satisfy (11), there is no guarantee that solving (11) gives a treatment confounding bridge function if multiple solutions exist. When a solution to (11) exists, we give the following assumptions that guarantee the uniqueness of solution.

Assumption 10 (Completeness).

(a) For any square-integrable function g, if $E[g(Z)|A, U, X] = 0$ almost surely, then $g(Z) = 0$ almost surely;

(b) For any square-integrable function h, if $E[h(U)|A, W, X, Y = 0, S = 1]$ almost surely, then $h(U) = 0$ almost surely.

Intuitively, Assumption 10 is a statement on the information contained in the unmeasured confounders vs. in the negative control variables – Assumption 10(a) requires the confounders U are informative enough about Z in the target population and Assumption 10(b) requires the NCO W is informative enough about U in the subgroup $Y = 0, S = 1$, so that no information is lost when taking the two conditional expectations. Completeness conditions similar to Assumption 10 were originally introduced by Lehmann and Scheffe to identify the so-called unbiase minimum risk estimator (Lehmann and Scheffé, 2012a,b). In econometrics and causal inference literature, completeness conditions have been employed to achieve identifiability for a variety of nonparametric or semiparametric models, such as instrumental variable regression (D'Haultfoeuille, 2011; Newey and Powell, 2003), measurement error models (Hu and Schennach, 2008), synthetic control (Shi et al., 2021), and previous works in negative control methods (Cui et al., 2020; Miao, Shi, and Tchetgen

Tchetgen, 2018; Ying et al., 2021). The completeness condition holds for a wide range of distributions. Newey and Powell (2003) and D'Haultfoeuille (2011) provided justifications in exponential families and discrete distributions with finite support. Andrews (2011) constructed a broad class of bivariate distributions that satisfy the completeness condition.

Assumption 10 have several immediate consequences:

Proposition 2. (a) Under Assumptions 4, 6 and 10, for any square integrable function g such that $E[g(Z)|A, W, X, S = 0, Y = 1] = 0$ almost surely, then $g(Z) = 0$ almost surely.

(b) Under Assumptions 10(a) and 1-5, the treatment confounding bridge function is unique. That is, if two square-integrable functions $q(A, Z, X)$ and $q_1(A, Z, X)$ satisfy

$$E[q(a, Z, x)|A = a, U = u, X = x] = E[q_1(a, Z, x)|A = a, U = u, X = x]$$
$$= \frac{1}{P(A = a|U = u, X = x)}$$

for all a, u, x almost surely, then $q(A, Z, X) = q_1(A, Z, X)$ almost surely.

(c) Under Assumptions 4, 6, 8 and 10, Equation (11) has a unique solution $q^*(A, Z, X)$.

We prove Proposition 2 below. Proposition 2 states that the completeness conditions in Assumption 10 and the definitions of the negative control variables lead to a third completeness condition that only involves the study data. Proposition 2 states the uniqueness of the treatment confounding bridge function, $q(A, Z, X)$, and the solution to (11), $q^*(A, Z, X)$. The function $q^*(A, Z, X)$ can therefore be identified from the study data alone and is a good approximation of $q(A, Z, X)$ by Theorem 2.

Proof. (a) Suppose a square-integrable function $g(Z)$ satisfies

$$E[g(Z)|A, W, X, Y = 0, S = 1] = 0 \quad \text{almost surely.}$$

The left-hand side equals

$$E\left\{E\left[g(Z)|U, A, W, X, Y = 0, S = 1\right] | A, W, X, Y = 0, S = 1\right\}$$

$$\overset{A.6}{=} E\left\{E\left[g(Z)|U, A, X, Y = 0, S = 1\right] | A, W, X, Y = 0, S = 1\right\}$$

$$\overset{A.4}{=} E\left\{E\left[g(Z)|U, A, X\right] | A, W, X, Y = 0, S = 1\right\}$$

By Assumption 10(b), $E\left[g(Z)|U, A, X\right] = 0$ almost surely. Then by Assumption 10(a), $g(Z) = 0$ almost surely.

(b) We have

$$E[q(A, Z, X) - q_1(A, Z, X)|A, U, X] = 0$$

almost surely. By Assumtion 10(a), we have

$$q(A, Z, X) - q_1(A, Z, X) = 0$$

almost surely, or $q(A, Z, X) = q_1(A, Z, X)$ almost surely.

(c) If two square-integrable functions $q^*(A, Z, X) = q_1^*(A, Z, X)$ satisfies

$$E[q^*(A, Z, X)|A, W, X, Y = 0, S = 1] = E[q_1^*(A, Z, X)|A, W, X, Y = 0, S = 1]$$

$$= \frac{1}{P(A|W, X, Y = 0, S = 1)}$$

almost surely, then

$$E\left\{[q^*(A, Z, X) - q_1^*(A, Z, X)]|A, W, X, Y = 0, S = 1\right\} = 0.$$

Under Assumptions 4, 6 and 10, Proposition 2(a) holds and therefore

$$q^*(A, Z, X) = q_1^*(A, Z, X)$$

almost surely.

$\square$

H   Proof of Corollary 2 and further discussion

We first prove equation (12)

$$E[m(W, A, X)q^*(A, Z, X) - m(W, 1, X) - m(W, 0, X)|Y = 0, S = 1]$$

$$= E\left\{m(W, A, X)E\left[q^*(A, Z, X)|W, A, X, Y = 0, S = 1\right] - m(W, 1, X) - m(W, 0, X)|Y = 0, S = 1\right\}$$

$$\overset{(11)}{=} E\left\{m(W, A, X)\frac{1}{P(A|W, X, Y = 0, S = 1)} - m(W, 1, X) - m(W, 0, X)|Y = 0, S = 1\right\}$$

$$= E\Big\{E\left[m(W, A, X)\frac{1}{P(A|W, X, Y = 0, S = 1)}|W, X, Y = 0, S = 1\right] -$$

$$m(W, 1, X) - m(W, 0, X)|Y = 0, S = 1\Big\}$$

$$= E\left\{m(W, 1, X) + m(W, 0, X) - m(W, 1, X) - m(W, 0, X)|Y = 0, S = 1\right\}$$

$$= 0$$

In fact, one can show that any regular and asymptotically normal estimator of τ that satisfies (11) has influence function of the form

$$IF(W, Z, A, X) = -\left\{E\left[\left.\frac{\partial q(A, Z, X; \tau)}{\partial \tau}\right|_{\tau=\tau_0} m(W, A, X)(1 - Y)\right|S = 1\right]\right\}^{-1}(1 - Y)$$
$$\left(m(W, A, X)q(A, Z, X) - m(W, 1, X) - m(W, 0, X)\right)$$

for an arbitrary function $m(W, A, X)$. Therefore, any regular and asymptotically normal estimator of τ corresponds to the solution of the estimating equation (13) for some function $m(W, A, X)$.

To prove this result, we see that for any parametric submodel that satisfies (11) and is

indexed by s such that the true distribution corresponds to $s = 0$, we have

$$E_s\left\{(1-Y)\left[q(A,Z,X;\tau_s) - \frac{1}{f_s(A|W,X,Y=0,S=1)}\right]m(W,A,X)\Big|S=1\right\} = 0$$

and thus

$$\partial E_s\left\{(1-Y)\left[q(A,Z,X;\tau_s) - \frac{1}{f_s(A|W,X,Y=0,S=1)}\right]m(W,A,X)\Big|S=1\right\}/\partial s\Big|_{s=0} = 0.$$

Note that

$$\frac{\partial}{\partial s}E_s[(1-Y)q(A,Z,X;\tau_s)m(W,A,X)|S=1]$$
$$=E\left[\frac{\partial q(A,Z,X;\tau)}{\partial \tau}\Big|_{\tau=\tau_0}m(W,A,X)(1-Y)\Big|S=1\right]\frac{\partial \tau_s}{\partial s}\Big|_{s=0} +$$
$$E[(1-Y)q(A,Z,X)m(W,A,X)S(O|S=1)|S=1]$$

and

$$\frac{\partial}{\partial s}E_s\left[\frac{(1-Y)m(W,A,X)}{f_s(A|W,X,Y=0,S=1)}\Big|S=1\right]$$

$$=E\left[-(1-Y)m(W,A,X)\frac{\frac{\partial}{\partial s}f_s(A|W,X,Y=0,S=1)}{f^2(A|W,X,Y=0,S=1)}\Big|S=1\right]+$$

$$E\left[\frac{m(W,A,X)(1-Y)}{f(A|W,X,Y=0,S=1)}S(W,A,X,Y=0|S=1)\Big|S=1\right]$$

$$=E\left[-\frac{m(W,A,X)(1-Y)}{f(A|W,X,Y=0,S=1)}S(A|W,X,Y=0,S=1)\Big|S=1\right]+$$

$$E\left[\frac{m(W,A,X)(1-Y)}{f(A|W,X,Y=0,S=1)}S(W,A,X,Y=0|S=1)\Big|S=1\right]$$

$$=E\left[\frac{m(W,A,X)(1-Y)}{f(A|W,X,Y=0,S=1)}S(W,X,Y=0|S=1)\Big|S=1\right]$$

$$=E\left[(1-Y)\{m(W,1,X)-m(W,0,X)\}S(W,X,Y=0|S=1)\Big|S=1\right]$$

$$=E\left[(1-Y)\{m(W,1,X)-m(W,0,X)\}S(O|S=1)\Big|S=1\right]$$

Rearranging the terms, we have

$$\frac{\partial\tau_s}{\partial s}\Big|_{s=0}=$$

$$E\left[-\left\{E\left[\frac{\partial q(A,Z,X;\tau)}{\partial\tau}\Big|_{\tau=\tau_0}m(W,A,X)(1-Y)\Big|S=1\right]\right\}^{-1}(1-Y)\right.$$

$$\left.\left(m(W,A,X)q(A,Z,X)-m(W,1,X)-m(W,0,X)\right)S(O|S=1)\Big|S=1\right]$$

I Discussion on the rare disease assumption 7

Through this section, we set $c(X) = 1$.

We first describe a crucial identity that links the log risk ratio β_0 and the treatment confounding bridge function $q(A, Z, X)$.

Lemma 2. *Under Assumptions 1, 2, 3 and 5, we have*

$$\beta_0 = \log\left(\frac{E[q(A, Z, X)I(A = 1, Y = 1)|S = 1]}{E[q(A, Z, X)I(A = 0, Y = 1)|S = 1]}\right) \tag{26}$$

Proof. The right hand side of (26) is

$$\log\left(\frac{E[q(A, Z, X)I(A = 1, Y = 1)|S = 1]}{E[q(A, Z, X)I(A = 0, Y = 1)|S = 1]}\right)$$
$$= \log\left(\frac{E[q(A, Z, X)|A = 1, Y = 1, S = 1]}{E[q(A, Z, X)|A = 0, Y = 1, S = 1]}\right) + \log\left(\frac{P(Y = 1, A = 1|S = 1)}{P(Y = 1, A = 0|S = 1)}\right)$$

Note that for $a = 0, 1$,

$$E[q(A, Z, X)|A = a, Y = 1, S = 1]$$

$$= E\left\{E[q(A, Z, X)|U, X, A = a, Y = 1, S = 1]|A = a, Y = 1, S = 1\right\}$$

$$\stackrel{A.2}{=} E\left\{E[q(A, Z, X)|U, X, A = a]|A = a, Y = 1, S = 1\right\}$$

$$\stackrel{A.5}{=} E\left\{\frac{1}{P(A = a|U, X)}\middle| A = a, Y = 1, S = 1\right\}$$

$$= \int \frac{1}{P(A = a|U = u, X = x)} f(u, x|A = a, Y = 1, S = 1) \, \mathrm{d}u \, \mathrm{d}x$$

$$= \int \frac{1}{P(A = a|U = u, X = x)} \times$$

$$\frac{P(A = a|U = u, X = x)P(Y = 1|A = a, U = u, X = x)P(S = 1|Y = 1, A = a, U = u, X = x)}{P(A = a, Y = 1, S = 1)} \, \mathrm{d}u \, \mathrm{d}x$$

$$\stackrel{A.2, A.3}{=} \int \frac{\exp(\beta_0 a)P(Y = 1|A = 0, U = u, X = x)P(S = 1|Y = 1, U = u, X = x)}{P(A = a, Y = 1, S = 1)} \, \mathrm{d}u \, \mathrm{d}x$$

$$= \frac{\exp(\beta_0 a)}{P(A = a, Y = 1, S = 1)} \int P(Y = 1|A = 0, U = u, X = x)P(S = 1|Y = 1, U = u, X = x) \, \mathrm{d}u \, \mathrm{d}x.$$

Therefore, the right-hand side of (26) equals

$$\beta_0 + \log\left(\frac{P(A = 0, Y = 1, S = 1)}{P(A = 1, Y = 1, S = 1)}\right) + \log\left(\frac{P(Y = 1, A = 1|S = 1)}{P(Y = 1, A = 0|S = 1)}\right)$$

$$= \beta_0 + \log\left(\frac{P(A = 0, Y = 1|S = 1)P(S = 1)}{P(A = 1, Y = 1|S = 1)P(S = 1)}\right) + \log\left(\frac{P(Y = 1, A = 1|S = 1)}{P(Y = 1, A = 0|S = 1)}\right)$$

$$= \beta_0$$

$\square$

Let q^* be the function that satisfies (11):

$$E[q^*(a, Z, X)|A = a, W, X, Y = 0, S = 1] = \frac{1}{P(A = a|U = u, X, Y = 0, S = 1)}.$$

We introduce an additional regularity condition:

Assumption 11 (Uniform continuity). *For any fixed positive square-integrable function $g(U)$ and a small positive number $0 < \eta < 1$, there exists some $0 < \gamma = \gamma(g, \eta) < 0$ such that $1 - \gamma < \dfrac{E[g_1(U)|A, W, X, Y = 0, S = 1]}{E[g(U)|A, W, X, Y = 0, S = 1]} < \dfrac{1}{1-\gamma}$ a.e. implies $1 - \eta < \dfrac{g_1(U)}{g(U)} < \dfrac{1}{1-\eta}$ a.e., where $g_1(U)$ is a positive square integrable function.*

Assumption 11 requires that the inverse mapping of $g \to E[g(U)|A, W, X, Y = 0, S = 1]$ is sufficiently smooth. By Theorem 2 and Assumption 8, we have

$$
\begin{aligned}
(1 - \delta)^3 &< \frac{E[q^*(A, Z, X)|A, W, X, Y = 0, S = 1]}{E[q(A, Z, X)|A, W, X, Y = 0, S = 1]} \\
&= \frac{E\{E[q^*(A, Z, X)|A, W, U, X, Y = 0, S = 1]|A, W, X, Y = 0, S = 1\}}{E\{E[q(A, Z, X)|A, W, U, X, Y = 0, S = 1]|A, W, X, Y = 0, S = 1\}} \\
&\stackrel{A.6}{=} \frac{E\{E[q^*(A, Z, X)|A, U, X]|A, W, X, Y = 0, S = 1\}}{E\{E[q(A, Z, X)|A, U, X]|A, W, X, Y = 0, S = 1\}} \\
&\stackrel{A.5}{=} \frac{E\{E[q^*(A, Z, X)|A, U, X]|A, W, X, Y = 0, S = 1\}}{E\{E[\frac{1}{P(A|U, X)}|A, U, X]|A, W, X, Y = 0, S = 1\}} \\
&< \frac{1}{(1 - \delta)^3}
\end{aligned}
$$

By Assumption 11, this implies

$$
\frac{1 - \eta(\delta)}{P(A|U, X)} < E[q^*(A, Z, X)|A, U, X] < \frac{1}{(1 - \eta(\delta))P(A|U, X)}. \tag{27}
$$

The constant $\eta(\delta)$ is determined by the smoothness of the inverse mapping of $g \mapsto E[g(U)|A, W, X, Y = 0, S = 1]$.

$$
E[q^*(A, Z, X)|U = u, X, Y = 0, S = 1] = \frac{1}{P(A = a|U = u, X, Y = 0, S = 1)} \tag{28}
$$

for almost all a and u. Let

$$
\begin{aligned}
\beta_0^* &= \log\left(\frac{E[q^*(A, Z, X)AY|S = 1]}{E[q^*(A, Z, X)(1 - A)Y|S = 1]}\right) \\
&= \log\left(\frac{E[q^*(A, Z, X)|A = 1, Y = 1, S = 1]}{E[q^*(A, Z, X)|A = 0, Y = 1, S = 1]}\right) + \log\left(\frac{P(Y = 1, A = 1|S = 1)}{P(Y = 1, A = 0|S = 1)}\right).
\end{aligned}
$$

Under mild regularity, the estimator $\widehat{\beta}$ is regular and asymptotically linear for β_0^*, and therefore

$$
\widehat{\beta} = \beta_0^* + O_p(1/\sqrt{n}) = \beta_0 + (\beta_0^* - \beta_0) + O_p(1/\sqrt{n}).
$$

It suffices to study

$$
\beta_0^* - \beta_0 = \log\left(\frac{E[q^*(A, Z, X)|A = 1, Y = 1, S = 1]}{E[q^*(A, Z, X)|A = 0, Y = 1, S = 1]}\right) - \log\left(\frac{E[q(A, Z, X)|A = 1, Y = 1, S = 1]}{E[q(A, Z, X)|A = 0, Y = 1, S = 1]}\right)
$$

Notice that

$$
\begin{aligned}
&E[q^*(A, Z, X)|A = a, Y = 1, S = 1] \\
&= E\{E[q^*(A, Z, X)|A = a, U, X, Y = 1, S = 1]|A = a, Y = 1, S = 1\} \\
&\overset{A.4}{=} E\{E[q^*(A, Z, X)|A = a, U, X]|A = a, Y = 1, S = 1\} \\
&\overset{A.4}{=} E\{E[q^*(A, Z, X)|A = a, U, X, Y = 0]|A = a, Y = 1, S = 1\} \\
&\overset{(27)}{<} \frac{1}{1 - \eta(\delta)} E\left\{\frac{1}{P(A = a|U, X)}\bigg|A = a, Y = 1, S = 1\right\} \\
&\overset{A.5}{=} \frac{1}{1 - \eta(\delta)} E\left\{E[q(A, Z, X)|A, U, X]\bigg|A = a, Y = 1, S = 1\right\} \\
&\overset{A.4}{=} \frac{1}{1 - \eta(\delta)} E\left\{E[q(A, Z, X)|A, U, X, Y = 1, S = 1]\bigg|A = a, Y = 1, S = 1\right\} \\
&= \frac{1}{1 - \eta(\delta)} E[q(A, Z, X)|A = a, Y = 1, S = 1]
\end{aligned}
$$

and similarly,

$$E[q^*(A, Z, X)|A = a, Y = 1, S = 1] > (1 - \eta(\delta))E[q(A, Z, X)|A = a, Y = 1, S = 1].$$

Therefore, we have

$$1 - \eta(\delta) < \frac{E[q^*(A, Z, X)|A = a, Y = 1, S = 1]}{E[q(A, Z, X)|A = a, Y = 1, S = 1]} < \frac{1}{1 - \eta(\delta)}.$$

We conclude that

$$2\log(1 - \eta(\delta)) = \log((1 - \epsilon)^2) < \beta_0^* - \beta_0 < \log\left(\frac{1}{(1 - \epsilon)^2}\right) = -2\log(1 - \epsilon)$$

and thus

$$|\widehat{\beta} - \beta_0| < -2\log(1 - \eta(\delta)) + O_p(\frac{1}{\sqrt{n}}).$$

J Regularity conditions and proof of Theorem 3

We denote τ_0^* as the true value of τ such that $q(A, Z, X; \tau_0^*) = q^*(A, Z, X)$. We will give the regularity conditions and proof that $(\widehat{\beta}, \widehat{\tau})$ is a regular and asymptotically linear estimator of (β_0^*, τ_0^*). Here β_0^* and q^* are the biased versions of β_0 and q defined in Appendix I, respectively, although the biases are negligible when the infection is rare. Following Appendix I, $\widehat{\beta}$ is also a regular and asymptotically linear estimator of β_0 if

$$\sup_{a,w,u,x} P(Y = 1 | A = a, W = w, U = u, X = x) < \delta_n,$$

Assumption 11 holds and

$$\log(1 - \eta(\delta_n)) = o_p(\frac{1}{\sqrt{n}}).$$

A set of regularity conditions are

R.1 The function $\tau \mapsto q(A, Z, X; \tau)$ is Lipschitz in a neighborhood of τ_0^*; that is, for every τ_1 and τ_2 in a neighborhood of τ_0 and a measurable function $\dot{q}(A, Z, X)$ with $E[\dot{q}(A, Z, X)] < \infty$, we have $\|q(A, Z, X; \tau_1) - q(A, Z, X; \tau_2)\| \le \dot{q}(A, Z, X)\|\tau_1 - \tau_2\|$;

R.2 $E[q(A, Z, X; \tau_0^*)^2] < \infty$ and $E[m(W, A, X)^2] < \infty$;

R.3 The function $\tau \mapsto q(A, Z, X; \tau)$ is differentiable at τ_0^*. The derivative matrix $\Omega(\beta_0^*, \tau^*)$ is nonsingular;

R.4 $\frac{1}{n}\sum_{i=1}^n G_i(\widehat{\beta}, \widehat{\tau}) = o_p(n^{-1/2})$ and $(\widehat{\beta}, \widehat{\tau}) \xrightarrow{p} (\beta_0^*, \tau_0^*)$.

The condition R.1 and the fact that $\beta \mapsto \exp(-\beta A)$ is Lipschitz in a neighborhood of β_0^* imply the function $(\beta, \tau) \to M_i(\beta, \tau)$ is Lipschitz in a neighborhood of (β_0^*, τ_0^*) for every i. The remaining proof follows Van der Vaart (2000) Theorem 5.21.

K Estimating conditional causal RR in the presence of effect modification by measured confounders

Algorithm 2 below describes a straightforward extension of Algorithm 1 to estimation the conditional vaccine effectiveness $VE(x) = 1 - \exp(\beta_0(x))$ under Assumption 9 and a parametric model $\beta_0(X; \alpha)$ indexed by a finite-dimensional parameter α.

Algorithm 2 Negative control method to estimate conditional vaccine effectiveness from a test-negative design

---

1: Identify the variables in the data according to Figure 1(c), in particular the NCEs and NCOs.
2: Estimate the treatment confounding bridge function by solving the equation (13) with a suitable parametric model $q^*(A, Z, X; \tau)$ and a user-specified function $m(W, A, X)$. Write $\widehat{\tau}$ as the resulting estimate of τ.
3: Estimate α by solving

$$\frac{1}{n} \sum_{i=1}^{n} (-1)^{1-A_i} c(X_i) q^*(A_i, Z_i, X_i; \widehat{\tau}) \exp(-\beta_0(X_i; \alpha) A_i) = 0 \qquad (29)$$

Denote the resulting estimate of α as $\widehat{\alpha}$. The estimated conditional vaccine effectiveness is
$$\widehat{VE}(x) = 1 - \exp(\beta_0(x; \widehat{\alpha})).$$

We describe the large-sample properties of the estimator $(\widehat{\alpha}, \widehat{\tau})$ in the theorem below.

Theorem 5 (Inference of $(\widehat{\alpha}, \widehat{\tau})$)**.** *Under Assumptions 1- 2, 4-5, 6, 9 and suitable regularity conditions listed at the end of this section, the estimator $(\widehat{\alpha}, \widehat{\tau})$ in Algorithm 2, or equivalently, the solution to the estimating equation $\frac{1}{n} \sum_{i=1}^{n} \tilde{G}_i(\alpha, \tau) = 0$ is regular and asymptotically linear with influence function*

$$\widetilde{IF}(\alpha, \tau) = - \left[\tilde{\Omega}(\alpha, \tau)^T \tilde{\Omega}(\alpha, \tau) \right]^{-1} \tilde{\Omega}(\alpha, \tau)^T \tilde{G}_i(\alpha, \tau),$$

where

$$\tilde{G}_i(\alpha,\tau) = \begin{pmatrix} (-1)^{1-A_i} c(X_i) q^*(A_i, Z_i, X_i; \tau) Y_i \exp(-\beta_0(X;\alpha) A_i) \\ (1-Y_i) \left[m(W_i, A_i, X_i) q^*(A_i, Z_i, X_i; \tau) - m(W_i, 1, X_i) - m(W_i, 0, X_i) \right] \end{pmatrix}$$

and

$$\tilde{\Omega}(\alpha,\tau) = \left(E\left[\frac{\partial \tilde{G}_i(\alpha,\tau)}{\partial \alpha^T} \right], E\left[\frac{\partial \tilde{G}_i(\alpha,\tau)}{\partial \tau^T} \right] \right).$$

Here $c(X)$ is a user-specified function of X with the sample dimension as α.

Suppose that in Algorithm 2, one specifies $\beta_0(X;\alpha) = X^T \alpha$, then a natural choice for $c(X)$ is $c(X) = X$. A sandwich estimator of the asymptotic variance of $(\widehat{\alpha}, \widehat{\tau})$ can be deduced from previous derivations. Under Assumption 9, we have shown that one can identify $VE(X)$, however, one may be unable to identify the population marginal VE without an additional assumption. Interestingly, we note that the population marginal risk ratio would remain non-identified even if one had access to a random sample from the target population to inform the marginal distribution of X. Specifically, as shown in Huitfeldt, Stensrud, and Suzuki (2019), the marginal $RR = E[Y(1)]/E[Y(0)]$ can be written as $RR = E[RR(X)|Y(0) = 1]$, i.e. the average risk ratio among subjects who would contract say Influenza had they possibly contrary to fact, not been vaccinated against Influenza. However, the distribution of X within the group $Y(a = 0) = 1$ cannot be identified in presence of unmeasured confounding, thus ruling out identification of the population marginal RR.

Define q^* as before and

$$\beta_0^*(x) = \log\left(\frac{E[q^*(A, Z, x) I(A = 1, Y = 1)|S = 1, X = x]}{E[q^*(A, Z, x) I(A = 0, Y = 1)|S = 1, X = x]} \right).$$

Denote α_0^* as the value of α such that $\beta_0(x; \alpha_0^*) = \beta_0^*(x)$. Below we give the set of regularity conditions such that $(\widehat{\alpha}, \widehat{\tau})$ is a regular and asymptotically linear estimator of (α_0^*, τ_0^*):

R'.1 The function $\tau \mapsto q(A, Z, X; \tau)$ is Lipschitz in a neighborhood of τ_0^* and $\alpha \mapsto \beta_0(X; \alpha)$ is Lipschitz in a neighborhood of α_0^*.

R'.2 $E[q(A, Z, X; \tau_0^*)^2] < \infty$, $E[m(W, A, X)^2] < \infty$, $E[C(X)^2] < \infty$ and $E[\exp(-2\beta_0(X; \alpha_0^*))] < \infty$. ;

R'.3 The function $\tau \mapsto q(A, Z, X; \tau)$ is differentiable at τ_0^* and $\alpha \mapsto \beta_0(X; \alpha)$ is differentiable at α_0^*. The derivative matrix $\tilde{\Omega}(\alpha_0^*, \tau^*)$ is nonsingular;

R'.4 $\frac{1}{n} \sum_{i=1}^{n} \tilde{G}_i(\widehat{\alpha}, \widehat{\tau}) = o_p(n^{-1/2})$ and $(\widehat{\alpha}, \widehat{\tau}) \xrightarrow{p} (\alpha_0^*, \tau_0^*)$.

L Proof of Theorem 1$'$

Lemma 3. *Under Assumptions 2$'$, 3$'$ and 6$'$, we have*

$$P(A = a, Y = y | U, W, X, S = 1)$$

$$= \frac{1}{\xi} P(A = a | Y = 0, U, W, X, S = 1) P(Y = y | A = 0, U, W, X, S = 1) \exp(\beta_0' a y)$$

where $\xi = \sum_{a^*, y^*} P(A = a^* | Y = 0, U, W, X, S = 1) P(Y = y^* | A = 0, U, W, X, S = 1) \exp(\beta_0' a^* y^*)$.

Proof. Note that

$$\frac{P(Y = 1 | A = 1, U, W, X, S = 1) P(Y = 0 | A = 0, U, W, X, S = 1)}{P(Y = 1 | A = 0, U, W, X, S = 1) P(Y = 0 | A = 1, U, W, X, S = 1)}$$

$$\overset{A.6'}{=} \frac{P(Y = 1 | A = 1, U, X, S = 1) P(Y = 0 | A = 0, U, X, S = 1)}{P(Y = 1 | A = 0, U, X, S = 1) P(Y = 0 | A = 1, U, X, S = 1)}$$

$$= \exp(\beta_0') \times \frac{P(S = 1 | Y = 1, A = 1, U, X)}{P(S = 1 | Y = 0, A = 1, U, X)} \times \frac{P(S = 1 | Y = 0, S = 0, U, X)}{P(S = 1 | Y = 1, S = 0, U, X)}$$

$$\overset{A.2'}{=} \exp(\beta_0') \times \exp(h(U, X)) \times \exp(-h(U, X))$$

$$= \exp(\beta_0')$$

The result follows after Chen (2003). □

To prove Theorem 1$'$, we need to show

$$E[(-1)^{1-A} c(X) q(A, Z, X) Y \exp(-\beta_a A) | S = 1] = 0$$

It suffices to prove that

$$E[(-1)^{1-A} q(A, Z, X) Y \exp(-\beta_a A) | U, X, S = 1] = 0.$$

The left-hand side is

$$\sum_{a,y} \int (-1)^{1-a} y \exp(-\beta_0' a) q(a,z,X) P(A=a,Y=y|U,X,S=1) f(z|U,X,A=a,Y=y,S=1) dz$$

$$\stackrel{L.3}{=} \sum_{a,y} \int (-1)^{1-a} y \exp(-\beta_0' a) q(a,z,X) \times$$

$$\frac{1}{\xi} P(A=a|Y=0,U,X,S=1) P(Y=y|A=0,U,X,S=1) \exp(\beta_0' ay) f(z|U,X,A=a,Y=y,S=1) dz$$

$$= \sum_a \int (-1)^{1-a} \exp(-\beta_0' a) q(a,z,X) \times$$

$$\frac{1}{c} P(A=a|Y=0,U,X,S=1) P(Y=1|A=0,U,X,S=1) \exp(\beta_0' a) f(z|U,X,A=a,Y=1,S=1) dz$$

$$= \sum_a \frac{(-1)^{1-a}}{\xi} P(A=a|Y=0,U,X,S=1) P(Y=1|A=0,U,X,S=1) \times$$

$$\int q(a,z,X) f(z|U,X,A=a,Y=1,S=1) dz$$

$$\stackrel{A.5'}{=} \sum_a \frac{(-1)^{1-a}}{\xi} P(A=a|Y=0,U,X,S=1) P(Y=1|A=0,U,X,S=1) \times$$

$$\frac{1}{P(A=a|Y=0,U,X,S=1)}$$

$$= \sum_a \frac{(-1)^{1-a}}{\xi} P(Y=1|A=0,U,X,S=1)$$

$$=0$$

M   Proof of Theorem 2$'$

$$E[\tilde{q}(a, Z, X)|A = a, W, X, Y = 0, S = 1]$$

$$=E\left\{E[\tilde{q}(a, Z, X)|A = a, U, W, X, Y = 0, S = 1]|A = a, W, X, Y = 0, S = 1\right\}$$

$$\overset{A.6'}{=}E\left\{E[\tilde{q}(a, Z, X)|A = a, U, X, Y = 0, S = 1]|A = a, W, X, Y = 0, S = 1\right\}$$

$$\overset{A.4'}{=}E\left\{E[\tilde{q}(a, Z, X)|A = a, U, X]|A = a, W, X, Y = 0, S = 1\right\}$$

$$\overset{A.5'}{=}E\left\{\frac{1}{P(A = a|U, X, Y = 0, S = 1)}|A = a, W, X, Y = 0, S = 1\right\}$$

$$=\int \frac{1}{P(A = a|U = u, X, Y = 0, S = 1)} f(u|A = a, W, X, Y = 0, S = 1) du$$

$$=\int \frac{1}{P(A = a|U = u, X, Y = 0, S = 1)} \frac{f(u|W, X, Y = 0, S = 1)P(A = a|U = u, W, X, Y = 0, S = 1)}{P(A = a|W, X, Y = 0, S = 1)} du$$

$$\overset{A.6'}{=}\int \frac{1}{P(A = a|U = u, X, Y = 0, S = 1)} \frac{f(u|W, X, Y = 0, S = 1)P(A = a|U = u, X, Y = 0, S = 1)}{P(A = a|W, X, Y = 0, S = 1)} du$$

$$=\frac{1}{P(A = a|W, X, Y = 0, S = 1)} \int f(u|W, X, Y = 0, S = 1) du$$

$$=\frac{1}{P(A = a|W, X, Y = 0, S = 1)}$$

N   Simulation setting with binary unmeasured confounder

We generate the data of a size $N = 7,000,000$ general population according to Figure 1, with the following distribution:

$$U \sim Bernoulli(p_U);$$

$$Z|U \sim Bernoulli(p_{0Z} + p_{UZ}U);$$

$$A|U \sim Bernoulli(p_{0A} + p_{UA}U);$$

$$Y|A,Y \sim Bernoulli(\exp(\eta_{0Y} + \beta_0 A + \eta_{UY}U));$$

$$W|U \sim Bernoulli(p_{0W} + p_{UW}U);$$

$$D|U \sim Bernoulli(p_{0D} + p_{UD}U);$$

$$S|Y,D,W,U \sim Bernoulli(\max(Y,D,W)(p_{YS} + p_{UYS}U)).$$

In the above data generating process, to mimic a test-negative design platform, we created a binary indicator D for flu-like diseases other than W. The study sample contains subjects with $S = 1$. The distribution of S indicates that only subjects with at least one of Y, D and W equal to one will be recruited into the study sample. We chose the parameters as in Table S1, which resulted in an average study sample size of between around 48,000 and around 52,000.

To obtain the true treatment confounding bridge function, by (6), we have

$$\sum_z q(a,z)f(z|u,a) = \sum_z q(a,z)f(a|u,z)f(z|u)/f(a|u) = 1/f(a|u)$$

Table S1: Parameter values of the data generating distribution in the simulation with a binary unmeasured confounder.

Parameter	Value	Parameter	Value
p_U	0.5	p_{0Z}	0.2
p_{UZ}	0.4	p_{0A}	0.2
p_{UA}	0.4	η_{0Y}	$\log(0.01)$
β_0	$\log(0.2), \log(0.5), \log(0.7),$ or 0	η_{UY}	$\log(0.5)$
p_{0W}	0.02	p_{UW}	0.02
p_{0D}	0.02	p_{UD}	-0.015
p_{YS}	0.1	p_{UYS}	0.4

and thus

$$\sum_z q(a,z) f(a|u,z) f(z|u)$$

$$= q(a,0)[p_{0A} + p_{UA}u]^a [1 - (p_{0A} + p_{UA}u)]^{1-a}[1 - (p_{0Z} + p_{UZ}u)] +$$

$$q(a,1)[p_{0A} + p_{UA}u]^a [1 - (p_{0A} + p_{UA}u)]^{1-a}(p_{0Z} + p_{UZ}u)$$

$$= 1$$

for each u, a. We obtain that

$$\begin{pmatrix} q(0,0) \\ q(0,1) \end{pmatrix} = \begin{pmatrix} [1-p_{0A}](1-p_{0Z}) & [1-p_{0A}]p_{0Z} \\ [1-(p_{0A}+p_{UA})][1-(p_{0Z}+p_{UZ})] & [1-(p_{0A}+p_{UA})](p_{0Z}+p_{UZ}) \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} q(1,0) \\ q(1,1) \end{pmatrix} = \begin{pmatrix} p_{0A}(1-p_{0Z}) & p_{0A}p_{0Z} \\ (p_{0A}+p_{UA})[1-(p_{0Z}+p_{UZ})] & (p_{0A}+p_{UA})(p_{0Z}+p_{UZ}) \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

O   Simulation setting with a continuous unmeasured confounder

In simulation studies with a continuous unmeasured confounder, we generate data for a general population of 7,000,000 individuals following the distribution below:

$$U \sim Uniform(0, 1);$$

$$X \sim Uniform(0, 1);$$

$$A|U, X \sim Binomial(\text{expit}(\mu_{0A} + \mu_{UA}U + \mu_{XA}X))$$

$$Z|A, U, X \sim N(\mu_{0Z} + \mu_{AZ}A + \mu_{XZ}X + \mu_{UZ}U, \sigma_Z^2);$$

$$W|U, X \sim N(\mu_{0W} + \mu_{XW}X + \mu_{UW}U, \sigma_W^2);$$

$$Y|A, U, X \sim Binomial(\text{expit}(\mu_{0Y} + \beta A + \mu_{UY}U + \mu_{XY}X) + \mu_{UXY}UX)$$

$$D|U, X \sim Binomial(\text{expit}(\mu_{0D} + \mu_{XD}X + \mu_{UD}U))$$

$$S|U, X, Y, D \sim \max(Y, D) \times Binomial(\text{expit}(\mu_{0S} + \mu_{XS}X + \mu_{US}U + \mu_{UXS}UX)).$$

We chose the parameter values according to Table S2, which results in an average study sample size of between around 43,000 and around 47,000.

Table S2: Parameter values of the data generating distribution in the simulation with a continuous unmeasured confounder.

Parameter	Value	Parameter	Value
μ_{0A}	-1	μ_{UA}	-1
μ_{XA}	0.25	μ_{0Z}	0
μ_{AZ}	0.25	μ_{XZ}	0.25
μ_{UZ}	4	σ_Z	0.25
μ_{0Y}	$\log(0.01)$	β_0	$\log(0.2), \log(0.5), \log(0.7),$ or 0
μ_{UY}	-2	μ_{XY}	-0.25
μ_{0W}	0	μ_{XW}	0.25
μ_{UW}	2	σ_W	0.25
μ_{0D}	$\log(0.01)$	μ_{XD}	0.25
μ_{UD}	-0.2	μ_{0S}	-1.4
μ_{XS}	0.5	μ_{US}	2
μ_{UXS}	1		

P Logistic regression as a naïve approximate estimator of log risk ratio, ignoring unmeasured confounders

Figure S1 shows a causal diagram of a test-negative design with no unmeasured confounders. Again we assumed selection into the study S is independent of the subjects' treatment status A, given the subjects' infection status Y and other covariates X.
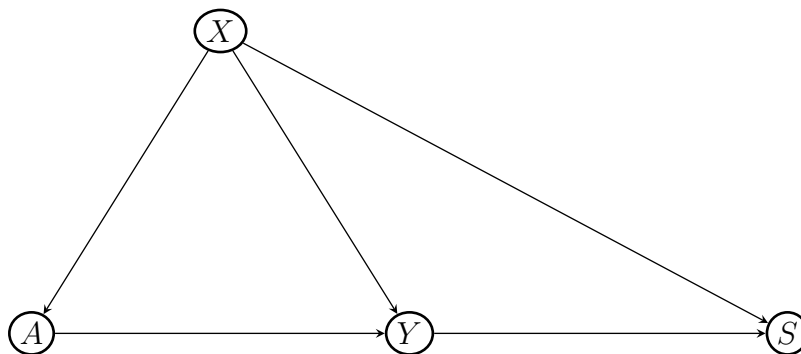

Figure S1: Directed acylic graph of a test-negative design with no unmeasured confounders.

In this scenario, we note that the conditional odds ratio given X in the study population

equals the conditional odds ratio in the general population, i.e.

$$\frac{P(Y=1|A=1,X,S=1)}{P(Y=1|A=0,X,S=1)} \bigg/ \frac{P(Y=0|A=1,X,S=1)}{P(Y=0|A=0,X,S=1)}$$

$$= \frac{P(Y=1|A=1,X)P(S=1|Y=1,A=1,X)}{P(Y=1|A=0,X)P(S=1|Y=1,A=0,X)} \bigg/ \frac{P(Y=0|A=1,X)P(S=1|Y=0,A=1,X)}{P(Y=0|A=0,X)P(S=1|Y=0,A=0,X)}$$

$$\stackrel{A.2}{=} \frac{P(Y=1|A=1,X)P(S=1|Y=1,X)}{P(Y=1|A=0,X)P(S=1|Y=1,X)} \bigg/ \frac{P(Y=0|A=1,X)P(S=1|Y=0,X)}{P(Y=0|A=0,X)P(S=1|Y=0,X)}$$

$$= \frac{P(Y=1|A=1,X)}{P(Y=1|A=0,X)} \bigg/ \frac{P(Y=0|A=1,X)}{P(Y=0|A=0,X)}$$

This implies that we may estimate the stratum-specific odds ratio in the general population by fitting a logistic regression model to the study data:

$$P(Y=1|A,X,S=1) = \text{expit}(\gamma_0 + \gamma_X X + \gamma_A A)$$

where $\text{expit}(x) = \exp(x)/[1 + \exp(x)]$, and γ_A is the log odds ratio. When the outcome is rare in the general population, i.e. $P(Y=0|A=a,X=x) \approx 1$ for any a, x, the log odds ratio γ_A also approximates the log risk ratio.

Q   Simulation for non-rare diseases

To investigate the performance of our method for non-rare diseases, we repeat the simulation with the same setup with binary or continuous confounders. For the simulation with binary confounders, we set η_{0Y} to be $\log(0.20)$; for the simulation with continuous confounders, we set μ_{0Y} to be $\log(0.20)$ – both correspond to an infection risk of 20% for subjects with $A = U = X = 0$.

The results are in Figure S2. While the NC-Oracle estimator remains unbiased and maintains calibrated confidence intervals, the NC estimator is in general biased with undercovered confidence intervals. Notably, the NC estimator is unbiased with calibrated confidence intervals under the null hypothesis where $\beta_0 = 0$. The logistic regression estimator is biased and has unpredictable behavior.
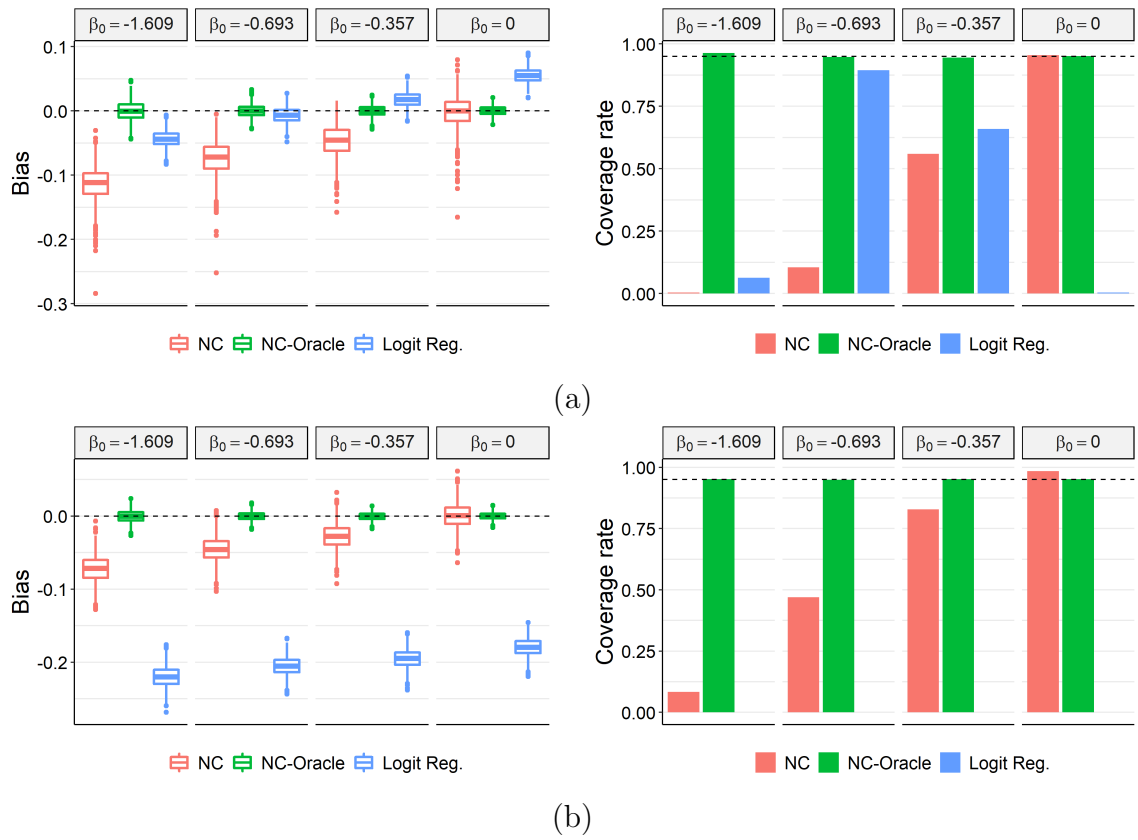
(a)

(b)

Figure S2: Bias (left) and coverage rates of 95% confidence interval (right) for the oracle estimator (NC-Oracle), GMM estimator (NC-GMM) and logistic regression (Logit Reg.) with a binary or unmeasured confounder, where η_{0y} or μ_{0Y} is $\log(0.20)$.

R Detailed results of University of Michigan Health System Data analysis

Table S3: Descriptive statistics of University of Michigan Health System COVID-19 Data. Variables were summarized as counts (percentage%).

	Unvaccinated (N=12,672)	Vaccinated (N=39,591)
Vaccine types		
Pfizer-BioNTech	/	20,312 (51.3%)
Moderna	/	10,831 (27.4%)
Johnson & Johnson's Janssen	/	1,409 (3.6%)
Other	/	7,039 (17.8%)
COVID-19 Infection	3,074 (24.2%)	2,774 (7.0%)
NCE: Immunization before Dec 2020	3,854 (30.4%)	18,167 (45.9%)
NCO conditions		
Arm/leg cellulitis	39 (0.3%)	161 (0.4%)
Eye/ear disorder	83 (0.6%)	518 (1.3%)
Gastro-esophageal disease	619 (4.9%)	3,188 (8.0%)
Atopic dermatitis	13 (0.1%)	41 (0.1%)
Injuries	1,033 (8.2%)	3,690 (9.3%)
General adult examination	752 (5.9%)	4,687 (11.8%)
No. of NCO conditions ≥ 1	2,258 (17.8%)	10,355 (26.2%)
Age		
≤ 18	1,683 (13.3%)	2,406 (6.1%)
$\geq 18, < 60$	8,667 (68.4%)	23,495 (59.3%)
≥ 60	2,322 (18.3%)	13,690 (34.6%)
Male	5,357 (42.3%)	16,241 (41.0%)
White	9.012 (71.1%)	30,748 (77.7%)
Charlson score ≥ 3	869 (6.8%)	3,230 (8.2%)

Table S4: Logistic regression of COVID-19 infection on COVID-19 vaccination, the NCE (previous immunization) and other baseline covariates.

	Est.	S.E.	p-value
(Intercept)	-1.30	0.06	<0.001
COVID-19 vaccination	-1.56	0.03	<0.001
Previous Immunization	0.37	0.03	<0.001
Age $\geq$18, $\leq$60	0.29	0.05	<0.001
Age $\geq$60	0.10	0.06	0.099
Male	0.01	0.03	0.750
White	0.32	0.04	<0.001
Charlson score $\geq$3	0.15	0.05	0.005
Calendar month			
April	-0.00	0.05	0.940
May	-0.98	0.06	<0.001
June	-1.79	0.08	<0.001
July	-1.16	0.07	<0.001
August	-0.53	0.05	<0.001
September	-0.40	0.05	<0.001
October	-0.10	0.04	0.026

Table S5: Logistic regression of having at least one NCO conditions on COVID-19 vaccination, the NCE (previous immunization) and other covariates.

	Est.	S.E.	p-value
(Intercept)	-2.32	0.06	<0.001
COVID-19 vaccination	0.28	0.03	<0.001
Previous Immunization	0.63	0.02	<0.001
Age $\geq$18, $\leq$60	0.28	0.04	<0.001
Age $\geq$60	0.76	0.05	<0.001
Male	-0.03	0.02	0.190
White	0.19	0.03	<0.001
Charlson score $\geq$3	0.38	0.04	<0.001
Calendar month			
April	-0.05	0.04	0.290
May	0.16	0.04	<0.001
June	0.16	0.04	<0.001
July	0.16	0.04	<0.001
August	0.14	0.04	<0.001
September	-0.03	0.04	0.360
October	0.04	0.04	0.309

References

Anders, Katherine L et al. (2018). "The AWED trial (Applying Wolbachia to Eliminate Dengue) to assess the efficacy of Wolbachia-infected mosquito deployments to reduce dengue incidence in Yogyakarta, Indonesia: study protocol for a cluster randomised controlled trial". In: *Trials* 19.1, pp. 1–16.

Andrews, Donald WK (2011). "Examples of l2-complete and boundedly-complete distributions". In.

Bareinboim, Elias and Judea Pearl (2012). "Controlling selection bias in causal inference". In: *Artificial Intelligence and Statistics*. PMLR, pp. 100–108.

Black, Carla L et al. (2018). "Influenza vaccination coverage among health care personnel—United States, 2017–18 influenza season". In: *Morbidity and Mortality Weekly Report* 67.38, p. 1050.

Bond, HS, SG Sullivan, and BJ Cowling (2016). "Regression approaches in the test-negative study design for assessment of influenza vaccine effectiveness". In: *Epidemiology & Infection* 144.8, pp. 1601–1611.

Boom, Julie A et al. (2010). "Effectiveness of pentavalent rotavirus vaccine in a large urban population in the United States". In: *Pediatrics* 125.2, e199–e207.

Broome, Claire V, Richard R Facklam, and David W Fraser (1980). "Pneumococcal disease after pneumococcal vaccination: an alternative method to estimate the efficacy of pneumococcal vaccine". In: *New England Journal of Medicine* 303.10, pp. 549–552.

Cai, Zhihong and Manabu Kuroki (2012). "On identifying total effects in the presence of latent variables and selection bias". In: *arXiv preprint arXiv:1206.3239*.

Chambers, Catharine et al. (2018). "Should sex be considered an effect modifier in the evaluation of influenza vaccine effectiveness?" In: *Open forum infectious diseases*. Vol. 5. 9. Oxford University Press US, ofy211.

Chen, Hua Yun (2003). "A note on the prospective analysis of outcome-dependent samples". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65.2, pp. 575–584.

Chung, Jessie R et al. (2020). "Effects of influenza vaccination in the United States during the 2018–2019 influenza season". In: *Clinical Infectious Diseases* 71.8, e368–e376.

Cole, Stephen R and Constantine E Frangakis (2009). "The consistency statement in causal inference: a definition or an assumption?" In: *Epidemiology* 20.1, pp. 3–5.

Cui, Yifan et al. (2020). "Semiparametric proximal causal inference". In: *arXiv preprint arXiv:2011.08411*.

Dagan, Noa et al. (2021). "BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting". In: *New England Journal of Medicine*.

Dean, Natalie E, Joseph W Hogan, and Mireille E Schnitzer (2021). *Covid-19 vaccine effectiveness and the test-negative design.*

Deaner, Ben (2018). "Proxy controls and panel data". In: *arXiv preprint arXiv:1810.00283.*

— (2021). "Many Proxy Controls". In: *arXiv preprint arXiv:2110.03973.*

Dufault, Suzanne M and Nicholas P Jewell (2020). "Analysis of counts for cluster randomized trials: Negative controls and test-negative designs". In: *Statistics in medicine* 39.10, pp. 1429–1439.

Dukes, Oliver, Ilya Shpitser, and Eric J Tchetgen Tchetgen (2021). "Proximal mediation analysis". In: *arXiv preprint arXiv:2109.11904.*

D'Haultfoeuille, Xavier (2011). "On the completeness condition in nonparametric instrumental problems". In: *Econometric Theory* 27.3, pp. 460–471.

Egami, Naoki and Eric J Tchetgen Tchetgen (2021). "Identification and Estimation of Causal Peer Effects Using Double Negative Controls for Unmeasured Network Confounding". In: *arXiv preprint arXiv:2109.01933.*

Feng, Shuo, Benjamin J Cowling, and Sheena G Sullivan (2016). "Influenza vaccine effectiveness by test-negative design–Comparison of inpatient and outpatient settings". In: *Vaccine* 34.14, pp. 1672–1679.

Fernández Villalobos, Nathalie Verónica et al. (2021). "Effect modification of the association between comorbidities and severe course of COVID-19 disease by age of study participants: a systematic review and meta-analysis". In: *Systematic reviews* 10.1, pp. 1–15.

Flannery, Brendan et al. (2019). "Influenza vaccine effectiveness in the United States during the 2016–2017 season". In: *Clinical Infectious Diseases* 68.11, pp. 1798–1806.

Foppa, Ivo M et al. (2016). "The case test-negative design for studies of the effectiveness of influenza vaccine in inpatient settings". In: *International journal of epidemiology* 45.6, pp. 2052–2059.

Gabriel, Erin E, Michael C Sachs, and Arvid Sjölander (2020). "Causal bounds for outcome-dependent sampling in observational studies". In: *Journal of the American Statistical Association*, pp. 1–12.

Ghassami, AmirEmad, Ilya Shpitser, and Eric Tchetgen Tchetgen (2021). "Proximal Causal Inference with Hidden Mediators: Front-Door and Related Mediation Problems". In: *arXiv preprint arXiv:2111.02927*.

— (2022). "Combining Experimental and Observational Data for Identification of Long-Term Causal Effects". In: *arXiv preprint arXiv:2201.10743*.

Ghassami, AmirEmad et al. (2021). *Minimax kernel machine learning for a class of doubly robust functionals.* Tech. rep.

Hitchings, Matt et al. (2021). "Effectiveness of the ChAdOx1 vaccine in the elderly during SARS-CoV-2 Gamma variant transmission in Brazil". In: *medRxiv*.

Hu, Yingyao and Susanne M Schennach (2008). "Instrumental variable treatment of non-classical measurement error models". In: *Econometrica* 76.1, pp. 195–216.

Hudgens, Michael G and M Elizabeth Halloran (2006). "Causal vaccine effects on binary postinfection outcomes". In: *Journal of the American Statistical Association* 101.473, pp. 51–64.

— (2008). "Toward causal inference with interference". In: *Journal of the American Statistical Association* 103.482, pp. 832–842.

Huitfeldt, Anders, Mats J Stensrud, and Etsuji Suzuki (2019). "On the collapsibility of measures of effect in the counterfactual causal framework". In: *Emerging themes in epidemiology* 16.1, pp. 1–5.

Imbens, Guido, Nathan Kallus, and Xiaojie Mao (2021). "Controlling for Unmeasured Confounding in Panel Data Using Minimal Bridge Functions: From Two-Way Fixed Effects to Factor Models". In: *arXiv preprint arXiv:2108.03849*.

Jackson, Lisa A et al. (2006). "Evidence of bias in estimates of influenza vaccine effectiveness in seniors". In: *International journal of epidemiology* 35.2, pp. 337–344.

Jackson, Michael L and Jennifer C Nelson (2013). "The test-negative design for estimating influenza vaccine effectiveness". In: *Vaccine* 31.17, pp. 2165–2168.

Jackson, Michael L et al. (2017). "Influenza vaccine effectiveness in the United States during the 2015–2016 season". In: *New England Journal of Medicine* 377.6, pp. 534–543.

Jewell, Nicholas P et al. (2019). "Analysis of cluster-randomized test-negative designs: cluster-level methods". In: *Biostatistics* 20.2, pp. 332–346.

Kallus, Nathan, Xiaojie Mao, and Masatoshi Uehara (2021). "Causal Inference Under Unmeasured Confounding With Negative Controls: A Minimax Learning Approach". In: *arXiv preprint arXiv:2103.14029*.

Krammer, Florian (2019). "The human antibody response to influenza A virus infection and vaccination". In: *Nature Reviews Immunology* 19.6, pp. 383–397.

Lehmann, Erich Leo and Henry Scheffé (2012a). "Completeness, similar regions, and unbiased estimation-Part I". In: *Selected Works of EL Lehmann*. Springer, pp. 233–268.

— (2012b). "Completeness, similar regions, and unbiased estimation—part II". In: *Selected Works of EL Lehmann*. Springer, pp. 269–286.

Leung, Jessica et al. (2011). "Herpes zoster incidence among insured persons in the United States, 1993–2006: evaluation of impact of varicella vaccination". In: *Clinical Infectious Diseases* 52.3, pp. 332–340.

Lipsitch, Marc, Ayan Jha, and Lone Simonsen (2016). "Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies". In: *International journal of epidemiology* 45.6, pp. 2060–2074.

Lipsitch, Marc, Eric Tchetgen Tchetgen, and Ted Cohen (2010). "Negative controls: a tool for detecting confounding and bias in observational studies". In: *Epidemiology (Cambridge, Mass.)* 21.3, p. 383.

Liu, Lan and Eric Tchetgen Tchetgen (2021). "Regression-based negative control of homophily in dyadic peer effect analysis". In: *Biometrics.*

Mastouri, Afsaneh et al. (2021). "Proximal causal learning with kernels: Two-stage estimation and moment restriction". In: *International Conference on Machine Learning.* PMLR, pp. 7512–7523.

Miao, Wang, Zhi Geng, and Eric J Tchetgen Tchetgen (2018). "Identifying causal effects with proxy variables of an unmeasured confounder". In: *Biometrika* 105.4, pp. 987–993.

Miao, Wang, Xu Shi, and Eric Tchetgen Tchetgen (2018). "A confounding bridge approach for double negative control inference on causal effects". In: *arXiv e-prints*, arXiv–1808.

Moline, Heidi L et al. (2021). "Effectiveness of COVID-19 vaccines in preventing hospitalization among adults aged≥ 65 years—COVID-NET, 13 states, February–April 2021". In: *Morbidity and Mortality Weekly Report* 70.32, p. 1088.

Newey, Whitney K and James L Powell (2003). "Instrumental variable estimation of nonparametric models". In: *Econometrica* 71.5, pp. 1565–1578.

Olson, Samantha M et al. (2022). "Effectiveness of BNT162b2 vaccine against critical Covid-19 in adolescents". In: *New England Journal of Medicine.*

Patel, Manish M, Michael L Jackson, and Jill Ferdinands (2020). "Postlicensure evaluation of COVID-19 vaccines". In: *JAMA* 324.19, pp. 1939–1940.

Qi, Zhengling, Rui Miao, and Xiaoke Zhang (2021). "Proximal Learning for Individualized Treatment Regimes Under Unmeasured Confounding". In: *arXiv preprint arXiv:2105.01187.*

Rolfes, Melissa A et al. (2019). "Effects of influenza vaccination in the United States during the 2017–2018 influenza season". In: *Clinical Infectious Diseases* 69.11, pp. 1845–1853.

Rosenbaum, Paul R (1987). "Model-based direct adjustment". In: *Journal of the American Statistical Association* 82.398, pp. 387–394.

Schnitzer, Mireille E (2022). "Estimands and Estimation of COVID-19 Vaccine Effectiveness Under the Test-negative Design: Connections to Causal Inference." In: *Epidemiology (Cambridge, Mass.)*

Schwartz, Lauren M et al. (2017). "Rotavirus vaccine effectiveness in low-income settings: An evaluation of the test-negative design". In: *Vaccine* 35.1, pp. 184–190.

Shi, Xu, Wang Miao, and Eric Tchetgen Tchetgen (2020). "A selective review of negative control methods in epidemiology". In: *Current Epidemiology Reports*, pp. 1–13.

Shi, Xu et al. (2020). "Multiply robust causal inference with double-negative control adjustment for categorical unmeasured confounding". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82.2, pp. 521–540.

Shi, Xu et al. (2021). "Theory for identification and Inference with Synthetic Controls: A Proximal Causal Inference Framework". In: *arXiv preprint arXiv:2108.13935.*

Shrank, William H, Amanda R Patrick, and M Alan Brookhart (2011). "Healthy user and related biases in observational studies of preventive interventions: a primer for physicians". In: *Journal of general internal medicine* 26.5, pp. 546–550.

Struchiner, CJ and ME Halloran (2007). "Randomization and baseline transmission in vaccine field trials". In: *Epidemiology & Infection* 135.2, pp. 181–194.

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". In: *American journal of epidemiology* 184.5, pp. 345–353.

Tchetgen Tchetgen, Eric J and Tyler J VanderWeele (2012). "On causal inference in the presence of interference". In: *Statistical methods in medical research* 21.1, pp. 55–75.

Tchetgen Tchetgen, Eric J et al. (2020). "An Introduction to Proximal Causal Learning". In: *arXiv preprint arXiv:2009.10982*.

Tenforde, Mark W et al. (2021). "Influenza vaccine effectiveness against hospitalization in the United States, 2019–2020". In: *The Journal of Infectious Diseases* 224.5, pp. 813–820.

Thompson, Mark G et al. (2021). "Effectiveness of COVID-19 vaccines in ambulatory and inpatient care settings". In: *New England Journal of Medicine* 385.15, pp. 1355–1371.

Utarini, Adi et al. (2021). "Efficacy of Wolbachia-infected mosquito deployments for the control of dengue". In: *New England Journal of Medicine* 384.23, pp. 2177–2186.

Van der Vaart, Aad. (2000). *Asymptotic statistics*. Vol. 3. Cambridge university press.

Wang, Bingkai et al. (2022). "Randomization Inference for Cluster-Randomized Test-Negative Designs with Application to Dengue Studies: Unbiased estimation, Partial compliance, and Stepped-wedge design". In: *arXiv preprint arXiv:2202.03379*.

Ying, Andrew et al. (2021). "Proximal Causal Inference for Complex Longitudinal Studies".

In: *arXiv preprint arXiv:2109.07030*.