medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Title page

Title: Potential Biases in Test-Negative Design Studies of COVID-19 Vaccine Effectiveness Arising from the Inclusion of Asymptomatic Individuals

Authors: Edgar Ortiz-Brizuela[1,2], Mabel Carabali[1], Cong Jiang[3], Joanna Merckx[1], Denis Talbot[4], *Mireille E. Schnitzer [3,5,1]

Affiliations:

1. Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada.
2. Department of Infectious Diseases, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico.
3. Faculty of Pharmacy, Université de Montréal, Montréal, Québec, Canada.
4. Department of Social and Preventive Medicine, Université Laval, Quebec, Quebec, Canada.
5. Department of Social and Preventive Medicine, School of Public Health, Université de Montréal, Montréal, Quebec, Canada.

Corresponding author: Mireille E. Schnitzer

Contact information:

Address: Faculty of Pharmacy, Université de Montréal, Montréal, QC, Canada.

Tel: 514 343-6111, extension: 0302

Email: mireille.schnitzer@umontreal.ca

Word Count:

Abstract: 195; Text: 3934.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

The test-negative design (TND) is a popular method for evaluating vaccine effectiveness (VE). A "classical" TND study includes symptomatic individuals tested for the disease targeted by the vaccine to estimate VE against symptomatic infection. However, recent applications of the TND have attempted to estimate VE against infection by including all tested individuals, regardless of their symptoms. In this article, we employ directed acyclic graphs and simulations to investigate potential biases in TND studies of COVID-19 VE arising from the use of this "alternative" approach, particularly when applied during periods of widespread testing. We show that the inclusion of asymptomatic individuals can potentially lead to collider stratification bias, uncontrolled confounding by health and healthcare-seeking behaviors (HSBs), and differential outcome misclassification. While our focus is on the COVID-19 setting, the issues discussed may also be relevant in the context of other infectious diseases. This may be particularly true in scenarios where there is either a high baseline prevalence of infection, a strong correlation between HSBs and vaccination, different testing practices for vaccinated and unvaccinated individuals, or settings where both the vaccine under study attenuates symptoms of infection and diagnostic accuracy is modified by the presence of symptoms.

Keywords

Test-negative design, COVID-19, vaccine, effectiveness, asymptomatic, SARS-CoV-2, simulation study, collider bias, confounding, misclassification.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

The post-licensure evaluation of COVID-19 vaccine effectiveness (VE), defined as the effect of vaccination on the risk of an infection-related outcome,[1] has provided crucial insights into questions not addressed through randomized clinical trials.[2-4] Among the study designs used to estimate VE, the test-negative design (TND) has gained popularity, in part because of its rapid and seemingly straightforward implementation.[2,5,6] In a "classical" TND study, symptomatic individuals tested for the disease targeted by the vaccine are prospectively selected, for example, from surveillance centers or hospitals.[5-8] Then, VE against symptomatic infection is estimated by comparing the odds of vaccination between patients with positive and negative test results using logistic regression.[5,7,8] However, recent literature[9-20] has also applied the term TND to studies aiming to estimate VE against infection by including all tested individuals, regardless of their symptom status (hereafter referred to as "alternative" TND) — for a clearer distinction between classical and alternative TND, see **Box 1**. Nonetheless, this approach may introduce additional threats to validity and warrants a more comprehensive evaluation.[21]

In this article, we investigate potential biases in TND studies of COVID-19 VE arising from the use of this "alternative" approach, particularly when applied during periods of widespread testing. We begin by discussing the identifiability of two causal target parameters: 1) the risk ratio (RR) for medically attended and laboratory-confirmed symptomatic COVID-19 relative to vaccination status, referred to as RR_{COVID} (the target parameter for the classical TND), and 2) the RR for SARS-CoV-2 infection relative to vaccination status, referred to as RR_{infect} (the target parameter for the alternative TND) — note that throughout this manuscript we use the expression "symptomatic COVID-19" to emphasize the difference between symptomatic and asymptomatic

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

infections. We conclude with a simulation study that aims to estimate and compare the magnitude of the bias in the odds ratio (OR) estimates for medically attended and laboratory-confirmed symptomatic COVID-19 (OR_{COVID}) and the OR for SARS-CoV-2 infection (OR_{infect}), relative to their respective target parameters (RR_{COVID} and RR_{infect}), under scenarios where a classical TND is considered valid.

Identifiability of target parameters

In this section, we use causal directed acyclic graphs (DAGs) to examine the identifiability of the target parameters (RR_{COVID} and RR_{infect}).[22] By "identifiability," we refer to the ability to express a counterfactual quantity as a function of the distribution of observed data, implying the absence of systematic biases.[23,24] However, before delving into the specifics of target parameter identifiability, and to facilitate a better understanding, we introduce the reader to a DAG based on previous work[7,8,21,25,26] that illustrates the assumed causal relationships between relevant variables in TND studies of COVID-19 VE, whether classical or alternative (**Box 2**).

The case of the classical TND

In this context, if the test-negative state (having some illness that presents with COVID-19-like symptoms – such as an infection not targeted by the vaccine ($I = 1$) – and testing) and the vaccination status are independent conditional on covariates, the case-status OR (OR_{COVID}) relative to vaccination status — derived from a correctly specified multivariable logistic regression model — provides an unbiased estimate of its target parameter: the conditional RR for medically attended and laboratory-confirmed symptomatic COVID-19 ($I = 2, S = 1, T = 1$) relative to vaccination status (RR_{COVID}).[8,27] Mathematical definitions for OR_{COVID} and RR_{COVID} are provided in **Box 3**. In

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

hospital-based TND studies, RR_{COVID} represents the risk ratio for medically attended infections leading to hospitalization (where T stands for "hospitalized and tested due to symptoms").[8] On the other hand, in outpatient-based TND studies, RR_{COVID} represents the risk ratio for medically attended infections in outpatient settings (where T stands for "accessed care and tested").

Importantly, unlike some traditional case-control studies, the TND does not require the rare disease assumption for the OR to approximate the RR.[8] It is also noteworthy that the OR_{COVID} estimate is often re-expressed as VE using the estimator: $\widehat{VE} = \left(1 - \widehat{RR}\right) \times 100$.[5] However, to avoid ambiguity, throughout this manuscript and unless otherwise specified, the term VE will continue to refer to any causal effect of vaccination on the risk of an infection-related outcome, regardless of the outcome assessed.[1]

In addition, the estimands mentioned above can be given a causal interpretation if the identifiability assumptions are also met (**Appendix S2**).[8,28,29] One such assumption, known as conditional exchangeability, implies that all common causes of vaccination status (V) and outcome (medically attended and laboratory-confirmed symptomatic SARS-CoV-2 infection $[I = 2, S = 1, T = 1]$) have been measured and appropriately accounted for in the analysis.[8,28-30] However, among these common causes, an individual's health and healthcare-seeking behaviors (HSBs) are inherently unmeasured.[21] In a best-case scenario, if HSBs were deterministic (meaning that individuals who exhibit these behaviors would always seek medical care when ill, while individuals who do not exhibit HSBs would not), a classical TND study could mitigate bias related to this variable by restricting the study population to individuals who have sought medical attention and are therefore assumed to exhibit HSBs ($H = 1$).[5,7] Although this scenario seems unlikely in real-world settings, where these behaviors may only modify the probability of seeking medical care, classical TND

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

studies could still offer better control for confounding by HSBs than other observational study designs such as cohorts or traditional case-control studies.[7,8]

To illustrate this point, **Figure 1(a)** shows a DAG that represents an ideal scenario for a classical TND study evaluating VE against medically attended and laboratory-confirmed symptomatic COVID-19. By "ideal", we mean that the study is conducted in a population where HSBs are deterministic and only perfect tests are used (with 100% sensitivity and specificity). In this study, the effect of interest is represented by the path $V \rightarrow I \bullet S \bullet T$. The square shapes of the nodes S, T, and H indicate that the study sample is restricted to symptomatic and tested individuals who sought medical care and therefore exhibit HSBs ($S = 1, T = 1, H = 1$). Conditioning on HSBs is critical because it not only allows control of confounding through the path $V \leftarrow H \rightarrow I$ but also blocks other biasing paths (e.g., $V \leftarrow H \rightarrow T \leftarrow S \leftarrow I$) that were opened by conditioning on colliders such as testing (T). In other words, restricting the study sample in a classical TND study to symptomatic and tested individuals is essential to prevent (or minimize) confounding and collider-stratification bias by HSBs,[7] and thus, to identify the RR_{COVID}.

The case of the alternative TND

Potential for collider stratification bias and uncontrolled confounding by HSBs

Having studied the classical TND, we now turn to the challenges presented by the alternative approach. In this scenario, since the study sample includes all tested individuals, regardless of symptoms, the target parameter is no longer the RR_{COVID}. Instead, it is the conditional RR for SARS-CoV-2 infection relative to vaccination status (RR_{infect}), represented by the path $V \rightarrow I$. Mathematical definitions for OR_{infect} and RR_{infect} are given in **Box 3**.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

To assess the identifiability of this target parameter, it is crucial to identify the specific drivers of testing leading to selection. Importantly, these drivers may have varied across settings, depending on the stage of the pandemic and the prevailing indications for testing.[31,32] For example, in a prospective, nonprobability-based, cross-sectional online survey of testing practices conducted between August 23, 2021, and March 12, 2022, among 418,279 U.S. adults aged ≥18 years, the most common reasons for testing, other than having symptoms, were exposure to COVID-19 (23.1%), a prerequisite for travel (20.6%), and requirement for work or school (16.1%).[31] Given that some of these reasons were mandatory at various points in time, we can no longer assume that the study sample is limited to healthcare seekers,[21] and that the TND effectively controls for bias related to HSBs. This may also be relevant to some hospital-based TND studies, as some institutions implemented policies for universal screening on admission.[33,34]

We illustrate this point in the DAG shown in **Figure 1(b)**, which represents an alternative TND study of VE against SARS-CoV-2 infection. In this DAG, the circular shapes of the nodes HSBs (H) and COVID-19-like symptoms (S) indicate that we are no longer conditioning on these variables. As a result, several biasing paths are opened, leading to uncontrolled confounding by HSBs ($V \leftarrow H \rightarrow I$; path 1 in **Figure S1**) and collider stratification bias ($I \leftarrow H \rightarrow T \leftarrow S \leftarrow V$ (path 2); $V \rightarrow S \leftarrow I$ (path 3); $I \leftarrow H \rightarrow T \leftarrow V$ (path 4); $I \rightarrow S \rightarrow T \leftarrow V$ (path 5); and $V \leftarrow H \rightarrow T \leftarrow S \leftarrow I$ (path 6)). In other words, the RR_{infect} is not identifiable in the alternative TND setting.

Potential for differential outcome misclassification

Another perceived benefit of the TND is its potential to reduce outcome misclassification compared to traditional cohort or case-control studies.[7,8] This is primarily attributed to the restriction of the study sample to individuals with confirmed infection status.[7,8] However, the

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

inclusion of asymptomatic individuals in alternative TND studies could potentially compromise this apparent benefit and instead lead to differential misclassification of the outcome relative to the exposure status. The main reason for this is that while SARS-CoV-2 diagnostic tests have excellent specificity for diagnosing acute infection, their sensitivity is likely modified by the presence of symptoms.[32] For example, two studies comparing the diagnostic accuracy of SARS-CoV-2 tests, reported antigen test sensitivities in symptomatic individuals of 72.0% (95% confidence interval [95% CI], 63.7% to 79.0%) and 58.1% (95% CI, 40.2% to 74.1%) in those asymptomatic,[35] and nucleic acid amplification tests (NAATs) sensitivities in symptomatic individuals of 97.1% (95% CI, 96.7% to 97.3%) and 87.6% (95% CI, 85.2% to 89.6%) in those without symptoms.[36]

This concept is illustrated in the DAG shown in **Figure 1(c)**. This DAG introduces a new node for the measured infection status (Ix), which may differ from its true value (I). There is differential misclassification of the outcome (I) relative to the exposure status (V) because, as described in **Box 2**, COVID-19 vaccines (V) may affect the presence of symptoms (S),[37-39] which in turn modify the sensitivity of SARS-CoV-2 tests (represented by the path $V \rightarrow S \rightarrow Ix$). This situation may be further complicated when considering that the selection of a specific test and its respective diagnostic performance, may depend on a number of additional factors.[32] These include the intended use of the test (i.e., screening versus confirmation, with molecular tests preferred for confirming diagnoses of acute infection in symptomatic individuals[32,40]), the time of infection onset,[41] the stage of the pandemic[32], and the target population (e.g., travelers vs. frontline workers[32]). As a result, it is difficult to accurately predict the magnitude and direction of this bias without a comprehensive understanding of the causal structure underlying a given study.[42]

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Simulation study

Methods

Overview

We conducted a simulation study to estimate and compare the magnitude of the bias in OR_{COVID} and OR_{infect} estimates — relative to their respective target parameters (RR_{COVID} and RR_{infect}) — under scenarios where a classical TND is considered valid. For simplicity, all simulations assume causal consistency,[30] no interference,[30] positivity,[30] and no unmeasured confounding other than that related to HSBs. To assess the bias pathways, we divided the simulation study into two parts. First, we evaluate the potential for collider bias and uncontrolled confounding by HSBs, assuming perfect diagnostic tests. Second, we incorporate the potential for differential outcome misclassification by simulating an extreme scenario where symptomatic individuals are tested exclusively with NAATs, and asymptomatic individuals are tested exclusively with antigen tests. Variables other than SARS-CoV-2 infection are assumed to be perfectly measured in all simulations.

Data generation process

To generate realistic data consistent with the DAG shown in **Box 2**, we generated 1,000 datasets (each with 1,000,000 observations) that simulated a cumulative prevalence of SARS-CoV-2 infection of ~12.2%,[43] including an asymptomatic infection prevalence of ~35%,[44] and a cumulative prevalence of fully vaccinated individuals of ~54.6%.[45,46] In addition, we simulated the proportion of individuals who reported using any SARS-CoV-2 diagnostic test (either NAATs or antigen tests) within 30 days in the aforementioned survey (~27.7%).[31] Each node in the DAG

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

was generated conditionally on its parent nodes using the models shown in **Table 1**, with parameters for these models selected based on the available literature whenever possible (**Table S1**). Importantly, although the simulations were designed to represent an outpatient setting, similar results can be expected for hospital-based TND studies, as the data generating structure may be similar for both scenarios (**Box 2**).

Data analysis

After generating the data, we first created classical and alternative TND samples from each simulated dataset by selecting either symptomatic and tested individuals or all tested individuals, respectively. Second, we estimated the OR_{COVID} (in the classical samples) and the OR_{infect} (in the alternative samples) using logistic regression models conditional on the set of measured confounders (C). For these models, we used the true SARS-CoV-2 infection status as the outcome variable when evaluating the potential for collider bias and uncontrolled confounding by HSBs, and the measured SARS-CoV-2 infection status, when incorporating the potential for differential outcome misclassification. Third, to estimate the true value of the target parameters (RR_{COVID} and RR_{infect}), we generated two counterfactual populations, each consisting of 100,000,000 individuals, representing scenarios in which everyone was either vaccinated or unvaccinated. Finally, we estimated the bias by subtracting the true values of the target parameters from the exponentiated mean of the log(OR) estimates obtained from the 1,000 simulated datasets (denoted as $\widehat{OR}_{COVID}$ or $\widehat{OR}_{infect}$). Additionally, we report the Monte Carlo standard error (MCSE, the standard deviation of the estimated log(ORs)) and the average standard error (aSE, the average of the standard error estimates of the log(ORs)).

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

To investigate the potential for larger biases, we iterated over the processes described above, each time varying the assumed strength of the relationships between the simulated variables. Specifically, we varied the magnitude of the association between H and V (parameter β_2 in **Table 1**), between H and both $I = 1$ and $I = 2$ (parameters $\gamma_{1,2}$ and $\gamma_{2,2}$, respectively), as well as the relationships of $I = 1$ and $I = 2$ with S (δ_4 and δ_5, respectively). We also varied the coefficient for the interaction term of $I = 2$ and V with S (δ_6), and the strengths of the relationships between H and T (θ_2), between V and T (θ_3), and between S and T (θ_4). We adjusted these parameters according to the direction of the association between the independent and dependent variables. Specifically, for parameters representing a positive association (i.e., OR > 1), the strength was increased by units of 1 from 1.5 to 10.5. Conversely, for parameters representing a negative association (i.e., OR < 1), the strength was decreased from 0.95 to 0.05 by units of 0.1. We also varied the intercepts for T (θ_0) and $I = 2$ ($\gamma_{2,0}$), aiming for a baseline prevalence of testing from 0.25 to 0.7 and for a baseline prevalence of acute SARS-CoV-2 infection from 0.05 to 0.5. Finally, we simultaneously varied the strength of two, and then three, of the most influential parameters (i.e., β_2, θ_3, and the intercept for $I = 2$). It should be noted that any adjustment to these parameters could lead to changes in the proposed prevalences for all simulated variables.

All analyses were performed using R (version 4.2.2, R Foundation for Statistical Computing, Vienna, Austria, 2022).[47]

Results

When we assessed the potential for collider bias and uncontrolled confounding by HSBs we found a "true value" for RR_{COVID} of 0.149 and an estimated $\widehat{OR}_{COVID}$ of 0.141 (MCSE and aSE $= 0.019$; bias $= -0.008$). On the other hand, the "true value" for RR_{infect} was 0.142, with an estimated

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

$\widehat{OR}_{infect}$ of 0.125 (MCSE and aSE = 0.014; bias = −0.016). The results after varying the strength of selected parameters for this scenario, either individually or simultaneously, are shown in **Figures 2-3** and **Tables S2-S3**. In every case, the classical TND outperformed the alternative in terms of bias, with the biases for $\widehat{OR}_{COVID}$ and $\widehat{OR}_{infect}$ ranging from −0.034 to -0.001 and from −0.211 to 0.728, respectively. When we incorporated imperfect tests, we found estimates for $\widehat{OR}_{COVID}$ of 0.154 (MCSE = 0.019; aSE = 0.018; bias = 0.005) and $\widehat{OR}_{infect}$ of 0.146 (MCSE and aSE = 0.014; bias = 0.004), respectively. However, after varying the strength of selected parameters for this scenario, either individually or simultaneously, we found larger bias for $\widehat{OR}_{infect}$ (ranging from −0.188 to 0.766) than for $\widehat{OR}_{COVID}$ (ranging from −0.005 to 0.018) (**Figures 4-5** and **Tables S4-S5**).

Discussion

The TND has become a preferred method for assessing VE against pathogens such as influenza, largely due to its seamless integration with existing laboratory-based surveillance systems.[6] More recently, this observational study design has been widely used to estimate VE against COVID-19.[21] This widespread application was made possible primarily by the extensive use of SARS-CoV-2 testing during the pandemic, which generated large amounts of potential data for TND studies.[21] However, the use of these data for VE estimation poses unique validity challenges.[21]

Originally, the TND was proposed as an efficient way to identify a control group (test-negative controls) for individuals who contracted the vaccine-targeted infection (test-positive cases) while also accounting for confounding by HSBs.[5,7,8] However, to effectively control for such bias in TND studies, researchers must assume that individuals seeking care and undergoing testing have

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

comparable HSBs.[5,7,8] In the pre-COVID-19 era, testing for diseases such as influenza was primarily driven by clinical indications and focused on those who actively sought care for symptoms.[48] This arguably made it easier for researchers to assume that individuals enrolled in TND studies had comparable HSBs, provided the same clinical definition was used for cases and controls. However, this paradigm shifted with COVID-19, when testing criteria were expanded to identify asymptomatic infections, often on a "mandatory" basis, in order to break transmission chains.[32] In this article, we have shown that TND applications that ignore this fact and include individuals regardless of their symptoms (i.e., using the alternative TND) can potentially introduce collider stratification bias, uncontrolled confounding by HSBs, and differential misclassification of the outcome relative to the exposure status.

Our simulation study, based on educated guesses for parameter values and representing a scenario in which the classical TND would yield unbiased estimates for its target parameter, showed only minor bias in the alternative TND setting. However, we also showed that the inclusion of asymptomatic individuals can lead to substantial bias in some scenarios. For example, our simulations showed that as the effect of HSBs on vaccination increases (OR > 1), the estimated $\widehat{OR}_{infect}$ can become progressively closer to the null relative to its target parameter, RR_{infect}, leading to further underestimation of VE when expressed as 1 – OR × 100. Similarly, in settings where vaccination status is more negatively correlated with testing (OR < 1), the $\widehat{OR}_{infect}$ would also be progressively biased towards the null. In addition, we observed that a higher baseline prevalence of SARS-CoV-2 infection would lead to a greater bias away from the null of $\widehat{OR}_{infect}$, resulting in an overestimation of VE. We also found that these biases may be exacerbated by the introduction of differential outcome misclassification when different diagnostic approaches are

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

used for symptomatic and asymptomatic individuals. Of note, the TND is well known to be particularly susceptible to misclassification bias compared to cohort and traditional case-control studies.[49] It is also important to mention that if one were to condition on the reason for testing or restrict the study sample to asymptomatic individuals, we would expect to observe the same bias described here, as all biasing pathways would remain open.

It could be argued that the biases discussed here may be only relevant to retrospective TND studies that rely on routinely collected health data (i.e., data collected without predetermined research questions[50]). For example, some TND studies using administrative data sources may have included asymptomatic individuals simply because of a lack of information on reasons for testing, as acknowledged by some authors.[14,15] However, it should be noted that some retrospective TND studies of COVID-19 VE intentionally included asymptomatic individuals to estimate VE against infection, even when data on symptoms were available.[10-12,16,17] In fact, some prospective TND studies (i.e., those collecting primary research data) conducted during the COVID-19 pandemic also included individuals regardless of symptoms.[18-20] That said, the potential for inclusion of asymptomatic individuals may be influenced not only by the data source, but also by the testing practices in the specific context or time frame in which the study is conducted.

The changing landscape of COVID-19 testing practices throughout the pandemic further highlights the need for proper attention to context in TND studies. At the onset of the COVID-19 pandemic, testing was primarily NAAT-based and focused on symptomatic individuals.[32] However, as the pandemic progressed, testing guidelines were modified to support mass testing of asymptomatic individuals as a tool for pandemic containment.[32] Any TND study conducted under these circumstances, whether prospective or retrospective, would be at risk of the bias described here if

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

a clinical definition was not included in the study's eligibility criteria. Thus, we believe that the issues discussed here are relevant not only to TND studies using routinely collected health and administrative data, but also to all studies conducted in settings where the rationale for testing extends beyond purely clinical reasons. **Box 4** provides a list of suggested strategies to minimize bias in TND studies conducted in such scenarios.

It is worth noting that there may be some specific circumstances in which the magnitude of the biases discussed here could be attenuated by other features of the study design, particularly those that allow for the selection of a population with comparable HSBs. For example, TND studies focusing on booster effectiveness may be less susceptible to HSBs-related bias because subjects in these studies typically have completed a primary immunization schedule and are therefore expected to have more similar HSBs.[21] Likewise, some might argue that hospital-based TND studies are likely to include individuals with more homogeneous HSBs because only those evaluated in clinical settings are eligible. However, as discussed elsewhere,[7] it may be unrealistic to assume that all hospitalized individuals have the same levels of HSBs. In addition, it should be noted that not all SARS-CoV-2 tests performed in hospitals under pandemic conditions were triggered by COVID-19-like symptoms; for example, some institutions mandated universal screening on admission.[33,34] Consequently, hospital-based TND studies that include all individuals tested regardless of the reason for testing,[13,51] would be expected to have biases similar to those discussed here.

This study has numerous strengths but also carries limitations. Among the strengths, our simulation study was illustrated by DAGs, which helps to clarify the potential sources of bias. In addition, while this article focuses on the COVID-19 setting, the issues discussed here are also relevant to

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

TND studies using the alternative approach in the context of other infectious diseases. This may be particularly true in scenarios of widespread testing, such as during epidemics or pandemics, where testing protocols and preferences are influenced by clinical and public health considerations, or in situations where self-testing for various infections is available.

In terms of limitations, the strength of the relationships depicted in the DAG may vary. For example, the path between vaccination and testing ($V \rightarrow T$) may be weak in some contexts. Nevertheless, collider bias could still occur through these nodes because unobservable factors, including HSBs, could confound the association between V and T. Another limitation lies in the values assigned to the parameters in our simulations. Although we relied on existing literature, these values are still estimates. However, varying these parameter values supported our theory-based hypothesis that relevant biases may occur in some settings. Finally, our study focused on specific issues related to the inclusion of asymptomatic individuals in TND studies. However, we did not assess additional complexities, such as the possibility of differential measurement error in exposure, outcome, or covariate status based on testing rationale, or the nuances in data quality and challenges specific to their respective sources, such as clinical versus other settings.[21]

Conclusions

In conclusion, the inclusion of asymptomatic individuals in TND studies of COVID-19 VE may lead to collider stratification bias, uncontrolled confounding by HSBs, and differential misclassification of the outcome relative to exposure status. Researchers designing or applying the TND for VE estimation need to be aware of these potential biases. Further research is needed to identify additional design or analytic strategies to control for biases related to HSBs that may improve the validity and utility of the TND for estimating VE against infection.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Data Availability Statement

Code and a sample simulated dataset are available and can be accessed here: https://ortizbrizuela.github.io/projects/

Acknowledgements

Work supported by the Canadian Institutes of Health Research Project Grant ECP-184178 (awarded to MES, DT, JM, CJ). MES holds the Canada Research Chair in Causal Inference and Machine Learning in Health Science. MC holds a Chercheur Boursier Junior 1 Award from the Fonds de recherche du Québec-Santé (FRQS). DT holds a Chercheur Boursier Junior 2 Award from FRQS. E-OB holds a Doctoral Training Award from FRQS.

Conflicts of interest:

None.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Cowling BJ, Sullivan SG. A concern over terminology in vaccine effectiveness studies. *Euro Surveill* 2018; **23**(10).
2. Patel MM, Jackson ML, Ferdinands J. Postlicensure Evaluation of COVID-19 Vaccines. *JAMA* 2020; **324**(19): 1939-40.
3. Prugger C, Spelsberg A, Keil U, Erviti J, Doshi P. Evaluating covid-19 vaccine efficacy and safety in the post-authorisation phase. *BMJ* 2021; **375**: e067570.
4. Barouch DH. Covid-19 Vaccines - Immunity, Variants, Boosters. *N Engl J Med* 2022; **387**(11): 1011-20.
5. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* 2013; **31**(17): 2165-8.
6. Chua H, Feng S, Lewnard JA, et al. The Use of Test-negative Controls to Monitor Vaccine Effectiveness: A Systematic Review of Methodology. *Epidemiology* 2020; **31**(1): 43-64.
7. Sullivan SG, Tchetgen Tchetgen EJ, Cowling BJ. Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness. *Am J Epidemiol* 2016; **184**(5): 345-53.
8. Schnitzer ME. Estimands and Estimation of COVID-19 Vaccine Effectiveness Under the Test-Negative Design: Connections to Causal Inference. *Epidemiology* 2022; **33**(3): 325-33.
9. Song S, Madewell ZJ, Liu M, Longini IM, Yang Y. Effectiveness of SARS-CoV-2 Vaccines against Omicron Infection and Severe Events: A Systematic Review and Meta-Analysis of Test-Negative Design Studies. *medRxiv* 2023: 2023.02.16.23286041.
10. Chemaitelly H, Tang P, Hasan MR, et al. Waning of BNT162b2 Vaccine Protection against SARS-CoV-2 Infection in Qatar. *N Engl J Med* 2021; **385**(24): e83.
11. Chemaitelly H, Yassine HM, Benslimane FM, et al. mRNA-1273 COVID-19 vaccine effectiveness against the B.1.1.7 and B.1.351 variants and severe COVID-19 disease in Qatar. *Nat Med* 2021; **27**(9): 1614-21.
12. Tang P, Hasan MR, Chemaitelly H, et al. BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. *Nat Med* 2021; **27**(12): 2136-43.
13. Bruxvoort KJ, Sy LS, Qian L, et al. Effectiveness of mRNA-1273 against delta, mu, and other emerging variants of SARS-CoV-2: test negative case-control study. *BMJ* 2021; **375**: e068848.
14. Skowronski DM, Setayeshgar S, Zou M, et al. Comparative Single-Dose mRNA and ChAdOx1 Vaccine Effectiveness Against Severe Acute Respiratory Syndrome Coronavirus 2, Including Variants of Concern: Test-Negative Design, British Columbia, Canada. *J Infect Dis* 2022; **226**(1): 485-96.
15. Ionescu IG, Skowronski DM, Sauvageau C, et al. BNT162b2 Effectiveness Against Delta and Omicron Variants of Severe Acute Respiratory Syndrome Coronavirus 2 in Adolescents Aged 12-17 Years, by Dosing Interval and Duration. *J Infect Dis* 2023; **227**(9): 1073-83.
16. Andrejko KL, Pry J, Myers JF, et al. Prevention of Coronavirus Disease 2019 (COVID-19) by mRNA-Based Vaccines Within the General Population of California. *Clin Infect Dis* 2022; **74**(8): 1382-9.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

17. Hitchings MDT, Ranzani OT, Torres MSS, et al. Effectiveness of CoronaVac among healthcare workers in the setting of high SARS-CoV-2 Gamma variant transmission in Manaus, Brazil: A test-negative case-control study. *Lancet Reg Health Am* 2021; **1**: 100025.
18. Li XN, Huang Y, Wang W, et al. Effectiveness of inactivated SARS-CoV-2 vaccines against the Delta variant infection in Guangzhou: a test-negative case-control real-world study. *Emerg Microbes Infect* 2021; **10**(1): 1751-9.
19. Singh C, Naik BN, Pandey S, et al. Effectiveness of COVID-19 vaccine in preventing infection and disease severity: a case-control study from an Eastern State of India. *Epidemiol Infect* 2021; **149**: e224.
20. Thiruvengadam R, Awasthi A, Medigeshi G, et al. Effectiveness of ChAdOx1 nCoV-19 vaccine against SARS-CoV-2 infection during the delta (B.1.617.2) variant surge in India: a test-negative, case-control study and a mechanistic study of post-vaccination immune responses. *Lancet Infect Dis* 2022; **22**(4): 473-82.
21. Shi X, Li KQ, Mukherjee B. Current Challenges With the Use of Test-Negative Designs for Modeling COVID-19 Vaccination and Outcomes. *Am J Epidemiol* 2023; **192**(3): 328-33.
22. Pearl J. Causal diagrams for empirical research. *Biometrika* 1995; **82**(4): 669-88.
23. Hernán MA RJ. Chapter 2. Randomized experiments. Causal Inference: What If: Chapman & Hall/CRC; 2020: 13-24.
24. Hernán MA RJ. Chapter 6. Graphical representation of causal effects. Causal Inference: What If: Chapman & Hall/CRC; 2020: 69-82.
25. Vandenbroucke JP, Pearce N. Test-Negative Designs: Differences and Commonalities with Other Case-Control Studies with "Other Patient" Controls. *Epidemiology* 2019; **30**(6): 838-44.
26. Ciocanea-Teodorescu I, Nason M, Sjolander A, Gabriel EE. Adjustment for Disease Severity in the Test-Negative Study Design. *Am J Epidemiol* 2021; **190**(9): 1882-9.
27. Dean NE, Hogan JW, Schnitzer ME. Covid-19 Vaccine Effectiveness and the Test-Negative Design. *N Engl J Med* 2021; **385**(15): 1431-3.
28. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias-a review of recent developments. *Int J Epidemiol* 2018; **47**(5): 1714-22.
29. Didelez V, Kreiner S, Keiding N. Graphical Models for Inference Under Outcome-Dependent Sampling. *Statistical Science* 2010; **25**(3): 368-87, 20.
30. Hernán MA RJ. Chapter 3. Observational studies. Causal Inference: What If: Chapman & Hall/CRC; 2020: 25-40.
31. Rader B, Gertz A, Iuliano AD, et al. Use of At-Home COVID-19 Tests - United States, August 23, 2021-March 12, 2022. *MMWR Morb Mortal Wkly Rep* 2022; **71**(13): 489-94.
32. Peeling RW, Heymann DL, Teo YY, Garcia PJ. Diagnostics for COVID-19: moving from pandemic response to control. *Lancet* 2022; **399**(10326): 757-68.
33. Stadler RN, Maurer L, Aguilar-Bultet L, et al. Systematic screening on admission for SARS-CoV-2 to detect asymptomatic infections. *Antimicrob Resist Infect Control* 2021; **10**(1): 44.
34. Talbot TR, Hayden MK, Yokoe DS, et al. Asymptomatic screening for severe acute respiratory coronavirus virus 2 (SARS-CoV-2) as an infection prevention measure in healthcare facilities: Challenges and considerations. *Infect Control Hosp Epidemiol* 2023; **44**(1): 2-7.
35. Dinnes J, Deeks JJ, Adriano A, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* 2020; **8**(8): CD013705.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

36. Hohl CM, Hau JP, Vaillancourt S, et al. Sensitivity and Diagnostic Yield of the First SARS-CoV-2 Nucleic Acid Amplification Test Performed for Patients Presenting to the Hospital. *JAMA Netw Open* 2022; **5**(10): e2236288.

37. Antonelli M, Penfold RS, Canas LDS, et al. SARS-CoV-2 infection following booster vaccination: illness and symptom profile in a prospective, observational community-based case-control study. *J Infect* 2023.

38. Antonelli M, Penfold RS, Merino J, et al. Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a prospective, community-based, nested, case-control study. *Lancet Infect Dis* 2022; **22**(1): 43-55.

39. Grana C, Ghosn L, Evrenoglou T, et al. Efficacy and safety of COVID-19 vaccines. *Cochrane Database Syst Rev* 2022; **12**(12): CD015477.

40. Government of Canada. (n.d.). Testing of vaccinated populations. Public Health Agency of Canada. Retrieved May 10, 2023, from https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/testing-screening-contact-tracing/testing-vaccinated-populations.html.

41. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med* 2020; **173**(4): 262-7.

42. Keogh RH, Shaw PA, Gustafson P, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1-Basic theory and simple methods of adjustment. *Stat Med* 2020; **39**(16): 2197-231.

43. World Health Organization. (n.d.). United States of America: WHO Coronavirus (COVID-19) Dashboard. Retrieved from (April 10, 2023): https://covid19.who.int/region/amro/country/us.

44. Sah P, Fitzpatrick MC, Zimmer CF, et al. Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. *Proc Natl Acad Sci U S A* 2021; **118**(34).

45. U.S. Census Bureau. (n.d.). U.S. and World Population Clock. United States Census Bureau. Retrieved from (April 10, 2023): https://www.census.gov/popclock/.

46. Mathieu E, Ritchie H, Ortiz-Ospina E, et al. A global database of COVID-19 vaccinations. *Nat Hum Behav* 2021; **5**(7): 947-53.

47. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

48. Uyeki TM, Bernstein HH, Bradley JS, et al. Clinical Practice Guidelines by the Infectious Diseases Society of America: 2018 Update on Diagnosis, Treatment, Chemoprophylaxis, and Institutional Outbreak Management of Seasonal Influenzaa. *Clin Infect Dis* 2019; **68**(6): e1-e47.

49. De Smedt T, Merrall E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. *PLoS One* 2018; **13**(6): e0199180.

50. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015; **12**(10): e1001885.

51. Tan CY, Chiew CJ, Pang D, et al. Vaccine effectiveness against Delta, Omicron BA.1, and BA.2 in a highly vaccinated Asian setting: a test-negative design study. *Clin Microbiol Infect* 2023; **29**(1): 101-6.

52. Wu F, Yuan Y, Li Y, et al. The acceptance of SARS-CoV-2 rapid antigen self-testing: A cross-sectional study in China. *J Med Virol* 2023; **95**(1): e28227.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

53. Centers for Disease Control and Prevention. (2021, April 2). CDC updates travel guidance for fully vaccinated people. Retrieved from (on April 13, 2023): https://www.cdc.gov/media/releases/2021/p0402-travel-guidance-vaccinated-people.html.

54. Occupational Safety and Health Administration. (n.d.). Employer Rights and Responsibilities Following a Federal OSHA Inspection. U.S. Department of Labor. Retrieved May 10, 2023, from https://www.osha.gov/sites/default/files/publications/OSHA4159.pdf.

55. World Health O. Evaluation of COVID-19 vaccine effectiveness: interim guidance, 17 March 2021. CC BY-NC-SA 3.0 IGO. Geneva: World Health Organization, 2021.

473 **Boxes**

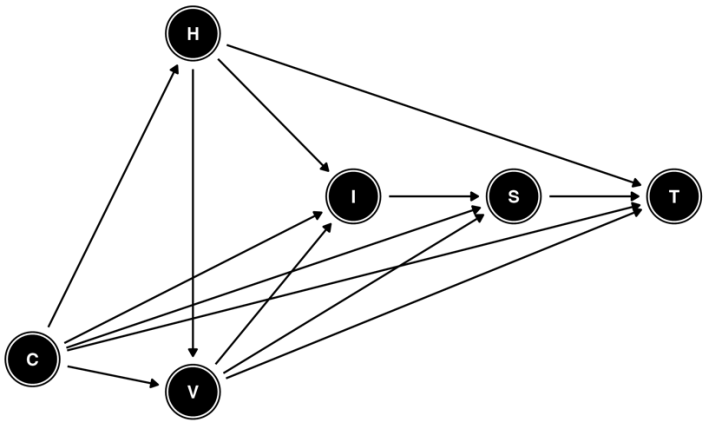
474 ***Box 1**. Relevant definitions used in this article.*

"Alternative" Test-Negative Design (TND)	An observational study design commonly used to estimate vaccine effectiveness (VE) against infection (whether symptomatic or asymptomatic). This design includes all individuals tested for the pathogen targeted by the vaccine under study, irrespective of their symptom status. VE is then estimated by comparing the odds of vaccination between individuals with positive and negative test results (test-positive cases and test-negative controls, respectively).
"Classical" TND	An observational study design commonly used to estimate VE against medically attended and laboratory-confirmed symptomatic infection. It includes individuals who are tested for the pathogen targeted by the vaccine under study due to the presence of symptoms. VE is then estimated by comparing the vaccination odds between individuals with positive and negative test results. Importantly, while both the alternative TND and the classical TND estimate VE, their specific objectives differ. While the alternative TND estimates VE against all infections, symptomatic or asymptomatic, the "classical" TND focuses only on VE against medically attended symptomatic infections. See below for definitions of infection and symptomatic infection.
Identifiability	The ability to express (identify) a counterfactual (causal) quantity, such as a measure of effect, as a function of the distribution of observed data, implying the absence of systematic bias.
Infection	The process by which a specific microorganism — the one targeted by the vaccine under study — invades and multiplies within an individual. An acute infection is usually confirmed by laboratory. Individuals with an infection may be asymptomatic, with no symptoms associated with the infection, or symptomatic, with symptoms associated with the infection. See the definition of “symptomatic infection” for more details.

Symptomatic Infection	This term describes an individual with a laboratory-confirmed infection (e.g., via nucleic acid amplification test) who also exhibit symptoms associated with that specific infection (e.g., fever, cough, or dyspnea in the case of COVID-19).
Target Parameter	The specific true measure of effect that a test-negative design study aims to estimate.
Vaccine Effectiveness	The effect of vaccination on the risk of an infection-related outcome. Specifically, it represents the expected reduction in the risk of an infection-related outcome that is directly attributable to vaccination.

475

***Box 2**. Directed acyclic graph for test-negative design studies of COVID-19 vaccine effectiveness (either classical or alternative).*



This causal Directed Acyclic Graph (DAG) consists of six nodes representing the following random variables:

1) V: Receipt of the COVID-19 vaccine under study (vaccinated $[V = 1]$, unvaccinated $[V = 0]$).
2) I: Infection status (no infection $[I = 0]$, an illness that presents with COVID-19-like symptoms but tests negative for COVID – such as a non-SARS-CoV-2 infection $[I = 1]$, a SARS-CoV-2 infection $[I = 2]$).
3) S: Presence of COVID-19-like symptoms (symptomatic $[S = 1]$, asymptomatic $[S = 0]$).
4) T: Whether a SARS-CoV-2 diagnostic test was performed (tested $[T = 1]$, not tested $[T = 0]$).
5) H: An indicator of the presence of health and healthcare-seeking behaviors (HSBs) that contribute to a greater probability of seeking medical care if COVID-19-like symptoms develop (HSBs present $[H = 1]$, HSBs absent $[H = 0]$).
6) C: A set of measured confounders of the association between vaccination status (V) and SARS-CoV-2 infection $(I = 2)$, such as age, calendar time, and sex.

Directed Acyclic Graph Rationale [* †]

The path $I \rightarrow S \rightarrow T$ represents the typical sequence of events leading to a COVID-19 diagnosis for a patient presenting with COVID-19-like symptoms. For example, an individual may contract a SARS-CoV-2 infection ($I = 2$) or another type of infection ($I = 1$), develop symptoms ($S = 1$), and subsequently seek care, resulting in COVID-19 testing ($T = 1$). Because this is an observational study, the association between COVID-19 vaccination (V) and SARS-CoV-2 infection ($I = 2$) is expected to be confounded by both the presence of HSBs (H) and the set of measured confounders(C).[7] Specifically, for this demonstration, we assume that individuals who exhibit HSBs ($H = 1$) are more likely to be vaccinated ($V = 1$) and to engage in other healthy or preventive behaviors that may protect them against SARS-CoV-2 infection ($I = 2$). Furthermore, certain individual characteristics, such as advanced age or the presence of comorbidities (i.e., the set of measured confounders C), may increase the probability that individuals will be vaccinated ($V = 1$) and the probability that they will become infected with SARS-CoV-2 ($I = 2$).

We also assume that an individual's HSBs (H) are influenced by the set of measured confounders (C).[7] In turn, these variables (H and C) influence an individual's probability of getting tested (represented by the path $H \rightarrow T \leftarrow C$). For example, older individuals, those with comorbidities, and those exhibiting HSBs ($H = 1$) may be more likely to get tested ($T = 1$) due to fear of potential complications.[52] Additionally, we assume that vaccinated individuals ($V = 1$) may be less likely to get tested ($T = 1$) in some situations. For example, vaccinated individuals ($V = 1$) may attribute their symptoms ($S = 1$) to alternative causes (e.g., allergies or other infections ($I = 1$)), or they may be exempt from certain testing policies (e.g., for some travel-related or workplace screening programs‡).[40,53,54] Moreover, we consider the possibility that vaccines may reduce symptoms or disease severity (e.g., hospitalization or death).[37-39] Therefore, we allow COVID-19-like symptoms (S) to be directly affected by COVID-19 vaccination (V). Finally, we consider that the set of measured confounders (C) may influence the presence of COVID-19-like symptoms (S) (represented by the path $C \rightarrow S$).[44]

Notes:

(*) For a brief introduction to the theory and terminology of DAGs, see **Appendix S1**.
(†) This DAG is applicable to both outpatient- and inpatient-based TND studies. In particular, the "testing" node can represent both hospitalization and testing in hospital-based TND studies (see Schnitzer, M. 2022 for an example).[8]
(‡) **Table S1** lists selected scenarios where testing practices varied between vaccinated and unvaccinated individuals during the pandemic.

477

478 ***Box 3**. Mathematical definitions for RR_{COVID}, OR_{COVID}, RR_{infect}, and OR_{infect}.*

Classical TND *

$$RR_{COVID} = \left(\frac{P(I = 2, S = 1, T = 1 \mid V = 1, H = 1, C)}{P(I = 2, S = 1, T = 1 \mid V = 0, H = 1, C)}\right)$$

$$OR_{COVID} = \left(\frac{P(I = 2 \mid V = 1, S = 1, T = 1, H = 1, C)}{P(I = 2 \mid V = 0, S = 1, T = 1, H = 1, C)}\right) / \left(\frac{P(I = 1 \mid V = 1, S = 1, T = 1, H = 1, C)}{P(I = 1 \mid V = 0, S = 1, T = 1, H = 1, C)}\right)$$

Alternative TND ‡

$$RR_{infect} = \left(\frac{P(I = 2 \mid V = 1, C)}{P(I = 2 \mid V = 0, C)}\right)$$

$$OR_{infect} = \left(\frac{P(I = 2 \mid V = 1, T = 1, C)}{P(I = 2 \mid V = 0, T = 1, C)}\right) / \left(\frac{P(I = 1 \mid V = 1, T = 1, C)}{P(I = 1 \mid V = 0, T = 1, C)}\right)$$

Notation:

- C: Confounders
- H: Health and healthcare-seeking behaviors
- $I = 1$: A test-negative COVID-19-like illness.
- $I = 2$: SARS-CoV-2 infection
- S: COVID-19-like symptoms
- T: SARS-CoV-2 diagnostic tests

- V: COVID-19 vaccination status

Notes:

(*) In the classical TND, the target parameter is the conditional risk ratio for medically attended and laboratory-confirmed symptomatic COVID-19. Therefore, the population of interest consists of individuals exhibiting HSBs ($H = 1$), with all SARS-CoV-2 infection ($I = 2$), symptoms ($S = 1$), and testing ($T = 1$) being part of the definition of RR_{COVID}. Furthermore, for OR_{COVID} to approximate RR_{COVID}, we require independence between the test-negative state (having some illness that presents with COVID-19-like symptoms – such as an infection non-targeted by the vaccine – and testing) and the vaccination status, conditional on covariates:

$$P(I = 1, S = 1, T = 1 \mid V = 1, H = 1, C) = P(I = 1, S = 1, T = 1 \mid V = 0, H = 1, C)$$

(‡) In the alternative TND setting, the target parameter is the conditional risk ratio for SARS-CoV-2 infection ($I = 2$), which includes both symptomatic and asymptomatic individuals. Since the study sample is not restricted to health care seekers, the target population is no longer limited to individuals with HSBs ($H = 1$). Our manuscript provides several reasons why this target parameter may not be identified under the alternative TND, i.e., why $OR_{infect} \neq RR_{infect}$.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

480 ***Box 4**. Suggested strategies for minimizing bias in TND studies conducted during infectious*

481 *disease outbreaks.*

- Researchers planning to assess vaccine effectiveness (VE) using the test-negative design (TND) should prioritize the validated – classical – form of this design and therefore must establish common eligibility criteria, including a clinical definition, for both test-positive cases and test-negative controls.
- To effectively implement the validated - classical - form of TND, investigators should collect comprehensive symptoms data prior to testing, including their type and time of onset, as well as data on other reasons for testing. This detailed information may allow investigators to apply tailored eligibility criteria, thereby facilitating comparability of cases and controls in terms of health and healthcare-seeking behaviors (HSBs).
- Because differences in the clinical presentation of vaccine-preventable infectious diseases may affect an individual's HSBs, researchers are advised not to pool data from inpatient and outpatient settings when using the TND.[6] This step is important to maintain comparability between cases and controls with respect to HSBs.
- When TND studies rely on routinely collected health and administrative data, investigators should attempt to obtain information on symptoms before testing (e.g., at admission for inpatient-based TND studies). If symptom information is not available, investigators may use alternative sources to define clinical eligibility criteria. For example, they may use hospital admission diagnoses as coded by the WHO International Classification of Diseases.
- Investigators should select the most sensitive and specific tests available (e.g., NAATs for COVID-19) to minimize the potential for differential misclassification of outcomes in TND studies. In addition, researchers should consider the potential variation in diagnostic performance due to factors such as the timing of symptom onset. For example, to reduce misclassification errors, especially for infections where viral load peaks near the end of the incubation period, it may be advisable to prioritize testing near the time of symptom onset.[55]
- Given the many potential biases associated with the use of the TND to assess VE against infection, we recommend against using this design for this particular outcome.

482

Tables

***Table 1**: Models used in the simulation study to generate data consistent with the directed acyclic graph shown in Box 2.*

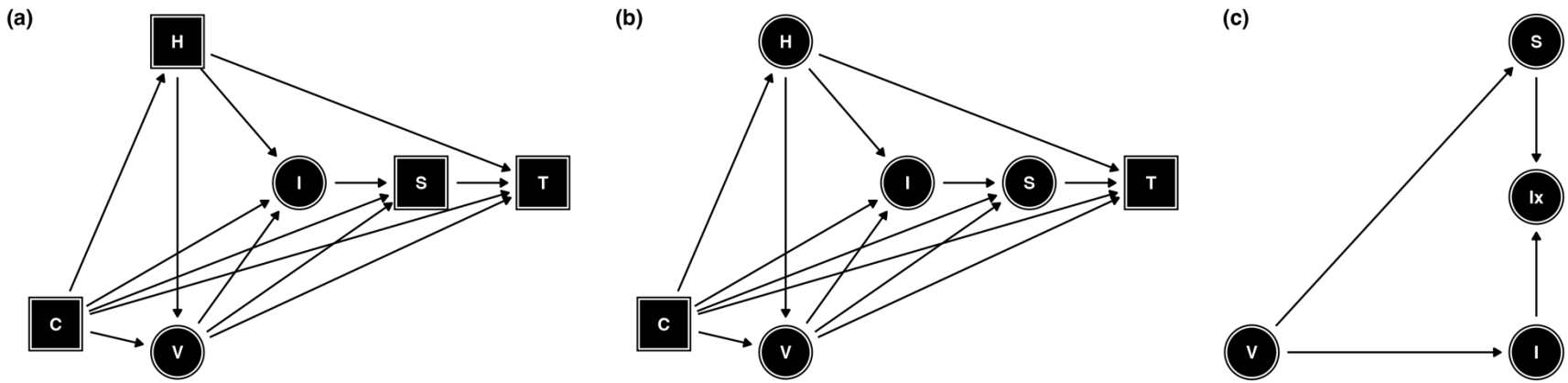
Variable (type)	Distribution / Parameters
***C*: Confounders (continuous)**	Gaussian; $\mu = 40, \sigma = 4$
***H*: Health and healthcare-seeking behaviors (binary: present vs. absent)**	Bernoulli; $\text{expit}(\alpha_0 + \alpha_1 C)$ where: • $\alpha_0 = -19.4$, and $\alpha_1 = \log(1.59)$. • For a prevalence of $H \sim 0.37$.
***V*: COVID-19 vaccination (binary: vaccinated vs. unvaccinated)**	Bernoulli; $\text{expit}(\beta_0 + \beta_1 C + \beta_2 H)$ where: • $\beta_0 = -12.25$, $\beta_1 = \log(1.36)$, and $\beta_2 = \log(1.81)$. • For a prevalence of $V \sim 0.55$.
***I* = 1: Non-SARS CoV-2 infection (binary: present vs. absent)**	Bernoulli; $\text{expit}(\gamma_{1,0} + \gamma_{1,1} C + \gamma_{1,2} H)$ where: • $\gamma_{1,0} = -6.7$, $\gamma_{1,1} = \log(1.12)$, and $\gamma_{1,2} = \log(0.9)$. • For a prevalence of $I = 1 \sim 0.10$.
***I* = 2: SARS-CoV-2 infection (binary: present vs. absent)**	Bernoulli; $\text{expit}(\gamma_{2,0} + \gamma_{2,1} C + \gamma_{2,2} H + \gamma_{2,3} V)*(1 - I[I = 1])$ where:

	• $\gamma_{2,0} = -5.4$, $\gamma_{2,1} = \log(1.12)$, $\gamma_{2,2} = \log(0.9)$, and $\gamma_{2,3} = \log(0.1)$. • For a prevalence of $I = 2 \sim 0.12$ and vaccine effectiveness of 90%. • The term (1- $I[I = 1]$) was incorporated under the assumption that $I = 1$ and $I = 2$ are mutually exclusive. • For simulations incorporating the potential for outcome misclassification, we generated the variable measured SARS-CoV-2 infection ($Ix = 2$) as shown in Figure 1(c), assuming a sensitivity of 0.97 for NAATs in symptomatic individuals and 0.58 for antigen tests in asymptomatic individuals, with a specificity of 0.99 for both tests.
***S*: COVID-19-like symptoms (binary: symptomatic vs. asymptomatic)**	Bernoulli; expit($\delta_0 + \delta_1 C + \delta_2 C{:}I[I = 1] + \delta_3 C{:}I[I = 2] + \delta_4 I[I = 1] + \delta_5 I[I = 2] + \delta_6 I[I = 2]{:}V) * (I[I = 1] + I[I = 2])$ where: • $\delta_0 = -33.8$, $\delta_1 = \log(1.2)$, $\delta_2 = \log(1.93)$, $\delta_3 = \log(1.93)$, $\delta_4 = \log(10)$, $\delta_5 = \log(10)$, and $\delta_6 = \log(0.52)$. • For a prevalence of $S \sim 0.65$ among those with SARS-CoV-2 infection ($I = 2$). • The term $(I[I = 1]) + I[I = 2])$ was included to ensure that symptoms only occur in individuals who have either $I = 1$ or $I = 2$.
***T*: COVID-19 diagnostic tests (binary: tested vs. not tested)**	Bernoulli; expit($\theta_0 + \theta_1 C + \theta_2 H + \theta_3 V + \theta_4 S) * (1 - I[H = 1, S = 1]) * (1 - I[H = 0, S = 1]) + I[H = 1, S = 1]$ where: • $\theta_0 = -38.8$, $\theta_1 = \log(2.47)$, $\theta_2 = \log(1.3)$, $\theta_3 = \log(0.92)$, and $\theta_4 = \log(6)$. • For a prevalence of $T \sim 0.28$.

	• The terms $(1 - I[H = 1, S = 1]) * (1 - I[H = 0, S = 1]) + I[H = 1, S = 1]$ were included to ensure that symptomatic individuals with HSBs would be tested while those without HSBs would not.

Figures

***Figure 1**: Causal-directed acyclic graphs for test-negative-design (TND) studies of COVID-19 vaccine effectiveness: a) the classical TND setting, b) the alternative TND setting, c) potential for differential misclassification of outcome status in alternative TND studies.*



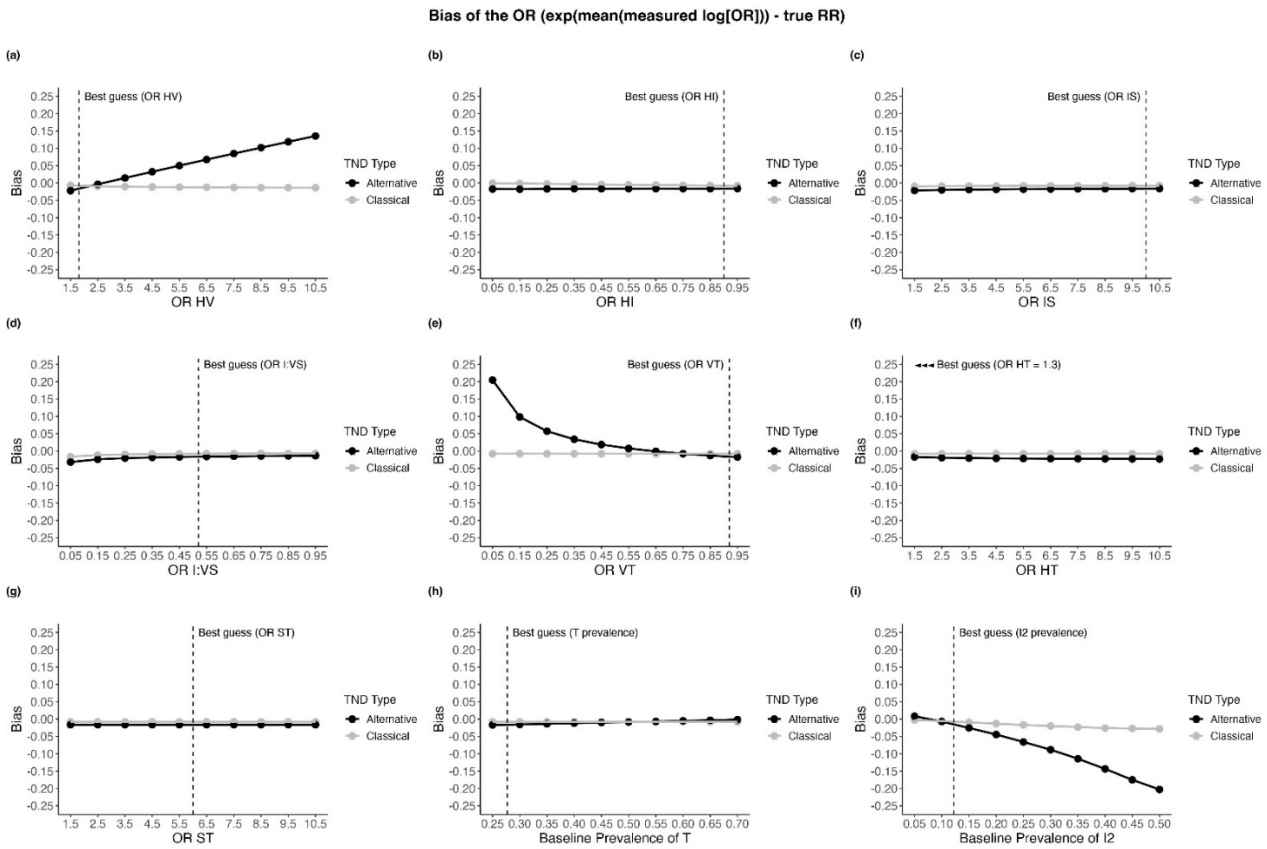
Abbreviations: C: Confounders; H: health and healthcare-seeking behaviors; I: Infection status; Ix: Measured infection status; S: COVID-19-like symptoms; T: SARS-CoV-2 diagnostic tests; V: COVID-19 vaccination status.

Notes: The square nodes indicate that the variable is controlled by either the study design or the analysis, while unconditioned nodes are represented as circles.

494 ***Figure 2**: Bias in odds ratio (OR) estimates for symptomatic COVID-19 (OR_{COVID}) and SARS-CoV-2 infection (OR_{infect}) relative to*

495 *their respective target parameters (RR_{COVID} and RR_{infect}) after varying the strength of selected parameters (assuming perfect SARS-*

496 *CoV-2 diagnostic tests).*

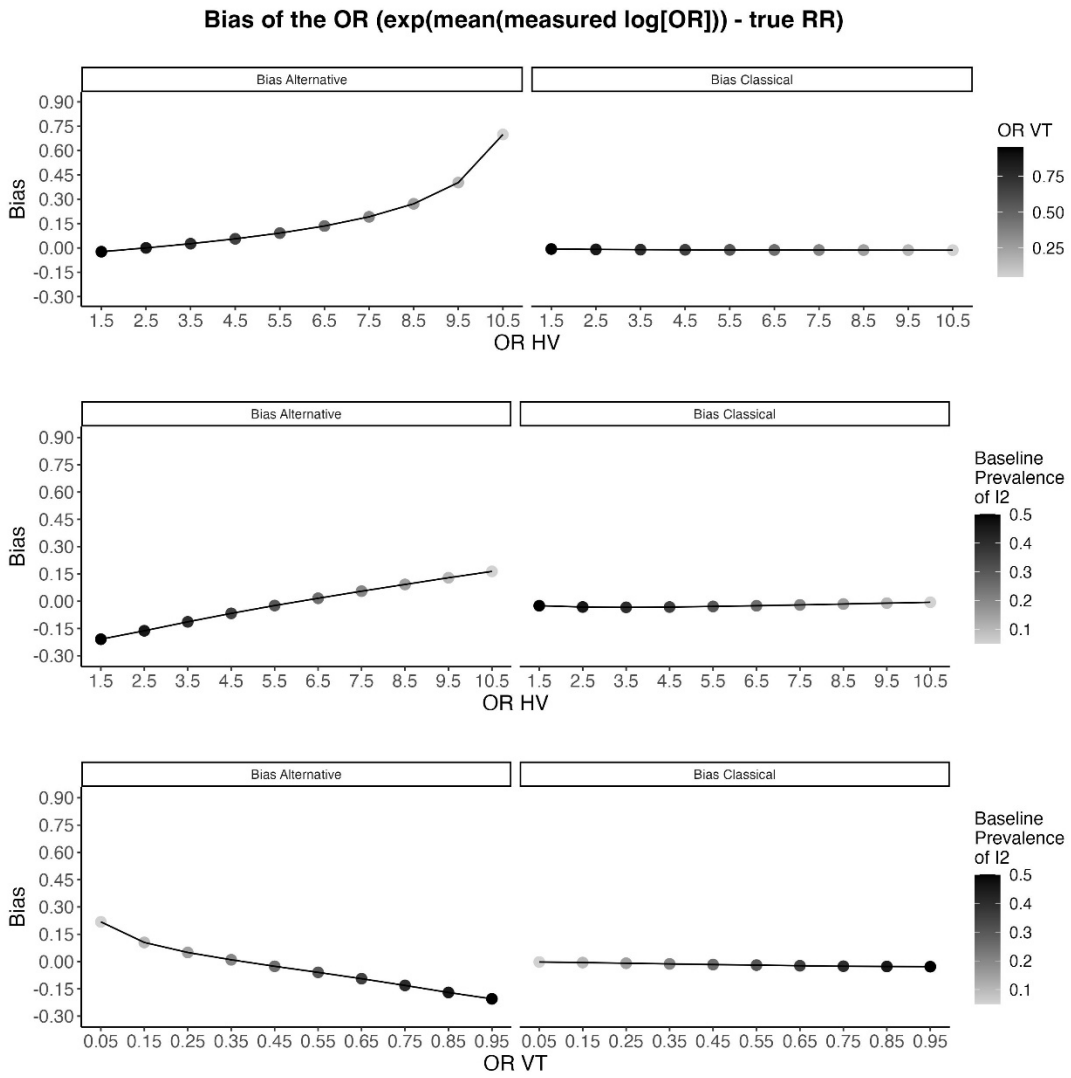


497

Abbreviations: C: Confounders; H: health and healthcare-seeking behaviors; I: Infection status; $I = 2$: SARS-CoV-2 infection status; OR, odds ratio; RR, risk ratio; S: COVID-19-like symptoms; T: SARS-CoV-2 diagnostic tests; TND, test-negative design; V: COVID-19 vaccination status.

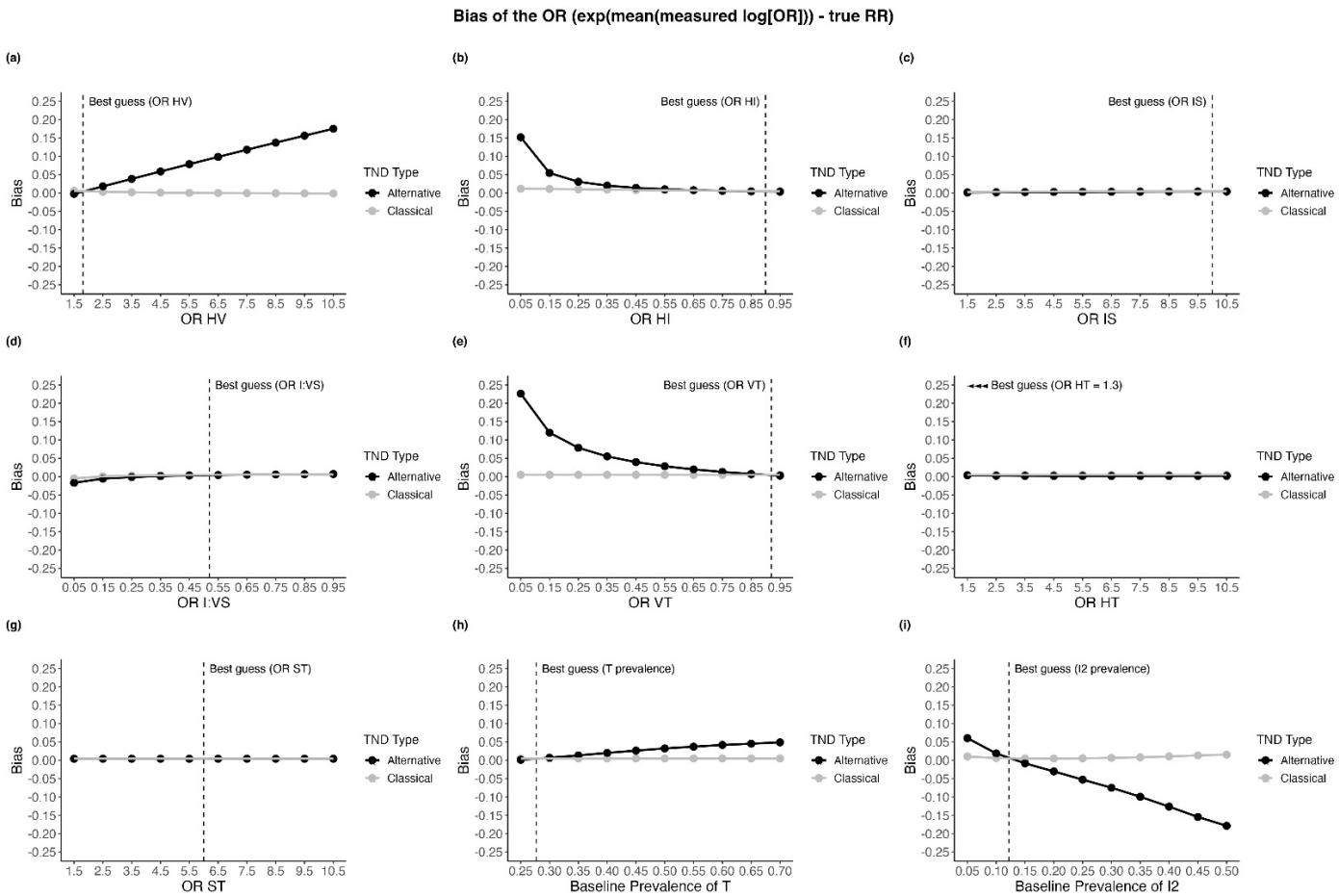
medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

***Figure 3**: Bias in odds ratio (OR) estimates for symptomatic COVID-19 (OR_{COVID}) and SARS-CoV-2 infection (OR_{infect}) relative to their respective target parameters (RR_{COVID} and RR_{infect}) after simultaneously varying the strength of two selected parameters (assuming perfect SARS-CoV-2 diagnostic tests).*



Abbreviations: H: health and healthcare-seeking behaviors; $I = 2$: SARS-CoV-2 infection status; NAATs, Nucleic Acid Amplification Tests; OR, odds ratio; RR, risk ratio; T: SARS-CoV-2 diagnostic tests; TND, test-negative design; V: COVID-19 vaccination status.

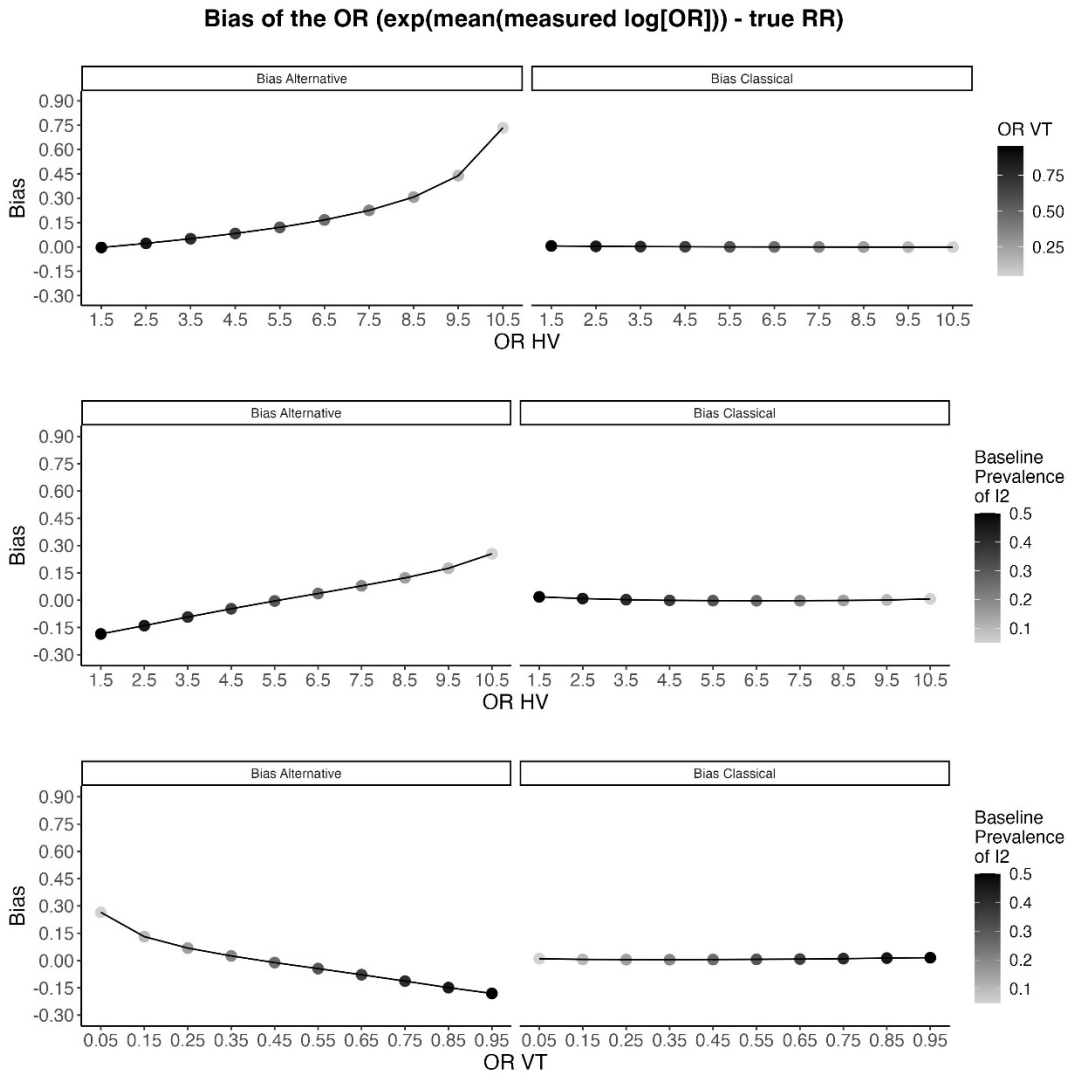
***Figure 4**: Bias in odds ratio (OR) estimates for symptomatic COVID-19 (OR_{COVID}) and SARS-CoV-2 infection (OR_{infect}) relative to their respective target parameters (RR_{COVID} and RR_{infect}) after varying the strength of selected parameters (assuming symptomatic individuals are tested exclusively with NAATs and asymptomatic individuals are tested exclusively with antigen tests).*



Abbreviations: C: Confounders; H: health and healthcare-seeking behaviors; I: Infection status; $I = 2$: SARS-CoV-2 infection status; NAATs, Nucleic Acid Amplification Tests; OR, odds ratio; RR, risk ratio; S: COVID-19-like symptoms; T: SARS-CoV-2 diagnostic tests; TND, test-negative design; V: COVID-19 vaccination status.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

***Figure 5**: Bias in odds ratio (OR) estimates for symptomatic COVID-19 (OR_{COVID}) and SARS-CoV-2 infection (OR_{infect}) relative to their respective target parameters (RR_{COVID} and RR_{infect}) after simultaneously varying the strength of two selected parameters (assuming symptomatic individuals are tested exclusively with NAATs and asymptomatic individuals are tested exclusively with antigen tests).*



Abbreviations: H: health and healthcare-seeking behaviors; $I = 2$: SARS-CoV-2 infection status; NAATs, Nucleic Acid Amplification Tests; OR, odds ratio; RR, risk ratio; T: SARS-CoV-2 diagnostic tests; TND, test-negative design; V: COVID-19 vaccination status.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298633; this version posted November 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

525 [---]