Commentary

Current Challenges With the Use of Test-Negative Designs for Modeling COVID-19 Vaccination and Outcomes

Xu Shi, Kendrick Qijun Li, and Bhramar Mukherjee*

* Correspondence to Dr. Bhramar Mukherjee, 1415 Washington Heights, Ann Arbor, MI 48109 (e-mail: bhramar@umich.edu).

The widespread testing for severe acute respiratory syndrome coronavirus 2 infection has facilitated the use of test-negative designs (TNDs) for modeling coronavirus disease 2019 (COVID-19) vaccination and outcomes. Despite the comprehensive literature on TND, the use of TND in COVID-19 studies is relatively new and calls for robust design and analysis to adapt to a rapidly changing and dynamically evolving pandemic and to account for changes in testing and reporting practices. In this commentary, we aim to draw the attention of researchers to COVID-specific challenges in using TND as we are analyzing data amassed over more than two years of the pandemic. We first review when and why TND works and general challenges in TND studies presented in the literature. We then discuss COVID-specific challenges which have not received adequate acknowledgment but may add to the risk of invalid conclusions in TND studies of COVID-19.

control selection; COVID-19; observational data; symptomatic testing; test-negative design; time-varying confounding; vaccine effectiveness

The test-negative design (TND) is increasingly used to evaluate postlicensure coronavirus disease 2019 (COVID-19) vaccine effectiveness (VE). In a TND study, patients who present at a health-care facility and get tested for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection are classified as cases (test-positive) and controls (test-negative). Then the odds of vaccination among cases and controls are compared to estimate VE against medically attended and laboratory-confirmed SARS-CoV-2 infection. VE against severe outcomes such as hospitalization and death may also be evaluated. The TND can be viewed as a type of case-control design, where the control subjects are drawn among those who tested negative. Besides estimation of VE, TND has also been used to study risk factors and other clinical outcomes of COVID-19 (1). Despite the comprehensive literature on TND, the use of TND in COVID-19 studies is relatively new and calls for robust design and analysis to adapt to a rapidly changing and dynamically evolving pandemic, with worldwide data amassed for more than 2 years. Over this time period, testing and reporting practices have also changed considerably.

WHEN AND WHY DOES TND WORK?

Figure 1A and 1B illustrate the rationale behind the TND and relationships among vaccination (V), true infection status (I), symptoms (S), testing (T), observed infection status (I^*), and health-care-seeking behavior (H), using directed acyclic graphs (DAGs). DAGs are widely used in epidemiology and econometrics to describe the variables under consideration (represented by nodes) and the causal relationships between them (represented by arrows). We highlight that testing can be due to clinical symptoms, in which case the outcome of interest is restricted to symptomatic infection, or testing could be due to other reasons unrelated to symptoms, in which case the outcome includes both symptomatic and asymptomatic infections. We further introduce the observed infection status I^*, which is not available among untested individuals. More discussions and alternative DAGs are available in other work (2–4). All arguments are implicitly conditional on the measured confounders such as age, gender, and socioeconomic status.

A major challenge TND aims to address is confounding by unmeasured health-care-seeking behavior when estimating

the vaccine effect against infection, indicated by the $V \rightarrow I$ arrow. Patients with higher levels of health-care-seeking behavior are more likely to get vaccinated ($H \rightarrow V$) and engage in healthy behaviors ($H \rightarrow I$). As such, vaccination may appear to be protective regardless of the true effect. The TND reduces such confounding bias by enrolling only the tested individuals, which is equivalent to conditioning the analysis on $T = 1$, leveraging the fact that patients with health-care-seeking behavior tend to seek care and get tested when ill ($H \rightarrow T$). We introduce the following assumptions:

- Assumption A: Tested patients have the same health-care-seeking behavior (i.e., $H = 1$ if $T = 1$).
- Assumption B: Vaccination status does not have a direct impact on testing (i.e., no $V \rightarrow T$ arrow).
- Assumption C: Vaccination does not mitigate the symptoms among individuals with the test-positive infection (i.e., no $V \rightarrow S$ arrow).

Under assumptions A, B, and C, the vaccinated and unvaccinated are comparable in terms of health-care-seeking behavior by design, which reduces confounding bias due to unmeasured health-care-seeking behavior in the estimated effect of vaccination.

In addition, when the outcome of interest is symptomatic infection, in traditional case-control studies, controls are randomly selected subjects, including those who may not seek testing/care and potentially have unreported COVID-19. That is, I^* is unknown if $T = 0$. In contrast, in a TND study where the outcome of interest is symptomatic infection, the control subjects are those who develop symptoms and seek care with SARS-Cov-2 test results available, and thus TND can reduce bias due to unreported cases when compared with case-control studies (5, 6). If the outcome includes both symptomatic and asymptomatic infection, misclassification bias due to asymptomatic testing is likely to remain and be similar under both designs.

CHALLENGES IN TND IN GENERAL

The TND is known to be subject to different sources of bias, which we illustrate in Figure 1C. First, as with all observational studies, TND may remain subject to unmeasured confounding bias (U). For example, occupation as a health-care worker, being residents of care facilities, needing a medical procedure, COVID-related comorbidities, and previous infection history are all potentially unmeasured common causes of vaccination, infection, and testing. Even when measured and included as an adjustment covariate, variables such as previous infection history are likely to be error-prone (3, 7). Circulating virus levels in the community may also affect both vaccination and testing. Adjusting for calendar time or daily test-positive rate can potentially control this bias. In addition, assumption A is almost never realistic because health-care-seeking behavior is multifaceted and cannot be fully adjusted by simply conditioning on testing.

Second, conditioning the analysis on testing may introduce collider stratification bias (8). A collider is a variable affected by at least 2 other variables, which, when conditioned on in analysis, can distort the association between the two variables. In Figure 1C, T is a collider of V and I due to the paths $V \rightarrow T$ and $I \rightarrow S \rightarrow T$, and conditioning on $T = 1$ leads to an observed association between V and I even when they are truly independent. For example, if unvaccinated individuals are more likely to get tested for SARS-CoV-2 infection due to, say, different SARS-CoV-2 testing policies by vaccination status at workplace, and individuals infected with SARS-CoV-2 are more likely to get tested as they develop symptoms, then a negative association between V and I may occur among the tested subjects. As another example, if the vaccine may mitigate symptoms due to SARS-Cov-2 infection, which leads to less testing, then T is a collider due to the $V \rightarrow S \rightarrow T$ and $I \rightarrow S \rightarrow T$ paths, which also leads to an observed negative association between V and I.

Other challenges discussed in the literature include: misclassification of infection status I^* due to imperfect testing (3, 6), lack of generalizability from the TND sample to the target population (9), confounding by calendar time due to time-varying vaccine uptake and infection acquisition (3, 5, 10), viral interference where infection by one virus affects infection by another virus (11), waning of VE (12), and differential susceptibility to infection due to prior infection (13). These issues have not received adequate acknowledgement, mostly due to lack of suitable data or statistical methods, but they add to the risk of invalid conclusions about VE.

CHALLENGES IN TND AS WE ANALYZE COVID-19 DATA

As we learn to live with COVID-19, we would like to draw the attention of researchers using TND to the following COVID-specific challenges in TND.

Comparing a booster dose with initial 2-dose vaccination

More recently, new TNDs have been implemented that allow researchers to compare vaccine-boosted individuals versus those who completed only primary series vaccination (14, 15). Several features of this comparison may be leveraged to improve the analysis and make more accurate causal conclusions. Because time of vaccination is available in both arms, it can be adjusted for in an analysis to remove time-window bias. Design strategies leveraging time of vaccination can be employed, such as using a negative control period of 14 days within the initial dose of the primary series vaccination, when no protection from vaccination is expected, to further detect and control for unmeasured confounding bias (16–18). Furthermore, as both arms have completed primary series vaccination, the VE estimate is less subject to confounding by health-care-seeking behavior.

Identifying covariates of relevance

Second, a growing fraction of the population has a history of prior infection. Adjustment for past SARS-CoV-2 infection history and differential test-positivity rate in geographical regions becomes important, as previous infection
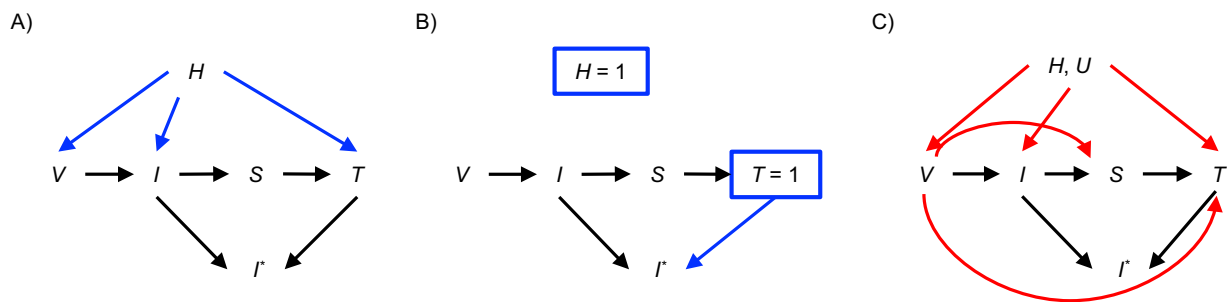
Figure 1. Directed acyclic graphs (DAGs) describing the causal relationship among vaccination status (*V*), true infection status (*I*), symptoms (*S*), testing (*T*), observed infection status (*I**), latent health-care-seeking behavior (*H*), and other unmeasured confounders (*U*). A and B) These DAGs illustrate how and why test-negative design works. C) This DAG presents scenarios when test-negative design can be subject to bias. If the outcome of interest is infection regardless of symptoms, the *H, U → T* arrow contains the effect through other reasons for testing.

and regional COVID-19 prevalence may be confounders between vaccination and new infection. It is also important to distinguish between natural immunity from infection and hybrid immunity from infection and vaccination by adjusting for prior SARS-CoV-2 infection. Moreover, attention should be paid when applying statistical methods that rely on rare-disease assumptions, which may be violated in certain study cohorts and as we move over time.

Test-negative design with widespread and regular testing

Third, the availability of testing is increasing. Individuals who self-test at home may not present at a health-care facility, which may affect our ability to recruit test-negative participants into a TND study. Alternative controls, such as nontested controls or population controls proposed to improve external validity (9, 19), may become less available. More and more tests conducted due to reasons other than health-care-seeking—such as traveling, work requirements, and medical procedures—will also affect the validity of TND to control for confounding bias (6, 20). With repeated testing becoming increasingly common, defining a case as someone with at least 1 reported positive test and a control as someone with all tests reported negative leads to individually varying sensitivity and specificity, depending on how many times one got tested. Existing TND studies handle the issue of repeated testing differently, for example, by limiting the maximum number of negative tests for someone with multiple negative tests or by randomly selecting a negative test result within specified intervals (21, 22). For someone with multiple illness periods with reinfections, typically only the first positive test in the initial period of illness is considered. All these protocols for choosing index tests are ad hoc. We need more systematic guidelines and research about the best practice for handling repeated testing in TNDs.

Lack of integrated public health data systems

Fourth, in order to implement a TND, we need to have a comprehensive database that captures all tests, vaccination, and other clinical outcomes for COVID-19 over time.

Such integrated public health and health-care systems are available in the United Kingdom, Israel, and Denmark (15, 17, 23), and as such seminal observations regarding waning effect of the vaccine and its reduced effectiveness against newer variants have emerged through studies using such large population-based data ecosystems. We need systems where each domain of data cross-talks with the other. Even to build a time-dependent model for "who is getting tested," we need individual-level data measured over time (24). Collecting data on the reason for getting tested helps us distinguish symptomatic testing.

Collecting accurate information on all-cause and cause-specific mortality and date of death by linking electronic health record data with state or national death registries may also help mitigate survivorship bias and competing risk bias (25). Survivor bias occurs since the study subjects in a test-negative design are necessarily alive by the time of testing, which restricts the conclusion of an analysis to the survivors. Competing risk bias occurs if death precludes the occurrence of testing and thus SARS-CoV-2 infection. Since COVID-19 vaccines appear to reduce mortality for the infected (26), among infected subjects who may die and who would otherwise get tested and be selected into the study sample, there may be an overrepresentation of unvaccinated subjects. This selection bias may cause an underestimation of the VE.

Time-varying confounding

Fifth, more than 2 years into the pandemic, we need to consider the time-varying features of confounding, vaccination, infection, and testing in the design of a study. For example, as illustrated in Figure 2, confounders of the vaccination-infection relationship, such as diagnosis of an autoimmune disease, can change over time; thus, timing of measurement for confounders should be carefully considered. Designs that only capture covariates between the index test and the last vaccination can miss important confounding information that influences both vaccination and infection status. Differential circulation of virus, transmissibility of emerging variants, change in policies, and human behaviors leading to time-varying infectiousness should also be accounted for.
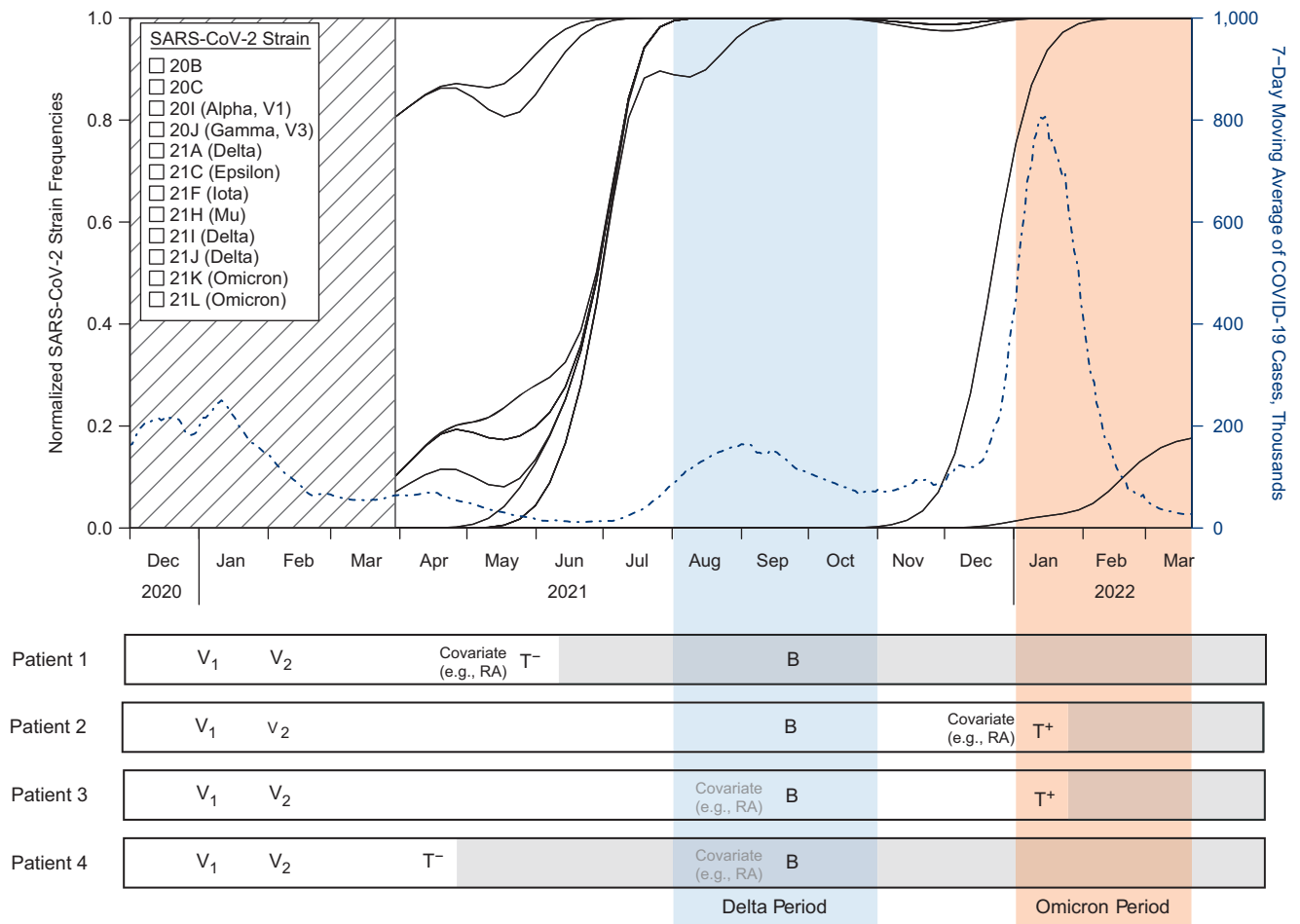
Figure 2. A schematic diagram depicting the evolving severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variant landscape (variant frequency data over time in the United States obtained from https://nextstrain.org), the reported coronavirus disease 2019 (COVID-19) case counts in the United States over time (obtained from the Covid Data Tracker (35)), and the possibility of time-varying confounding across some possible patient scenarios of vaccination and infection. V_1, V_2, and B represent the first, second, and booster vaccination; T^+ and T^- indicate test positive and negative. The colored shading indicates the Delta (blue) and Omicron (orange) variant periods. Patient 1 vs. 2: Differential circulation of virus and time-varying infectiousness should be accounted for. Patient 2 vs. 3: Measuring covariates between last vaccination and the index test may miss important confounders. Patient 3 vs. 4: Everything after the index test does not count in the comparison. RA, rheumatoid arthritis.

New designs and analysis

Finally, we need to be open to other forms of designs and analyses that may better address specific questions related to COVID-19. Some potential study designs are: using historic controls, using contemporaneous controls, self-controlled case series, and case-crossover studies. For example, to evaluate post-COVID new symptoms and health-care utilization, a case-crossover design (27), choosing proper windows of comparison during case (post-COVID) and control (pre-COVID) periods, may be a prudent choice.

Assessing VE against severe COVID-19 outcomes—such as hospitalization, intensive care unit admission, and death—is critical to inform decision making, particularly when clinical trials are inadequately powered for such rare outcomes. Some TND studies leverage hospital-based surveillance programs (28–30), in which patients admitted to participating hospitals meeting eligibility criteria are enrolled with infection and vaccination status prior to hospitalization being verified (i.e., data are collected prospectively) (3). Among hospitalized and tested patients, VE against both COVID-19 infection and severe COVID-19 outcomes among test-positive patients are studied. A distinction regarding being hospitalized "with COVID" and "for COVID" can be made in defining the hospitalization outcome. Other TNDs identify tested patients from electronic health records or other administrative databases with data collected retrospectively, and then investigate severe outcome among only test-positive patients (31, 32).

Besides using other study designs, we may also explore the utility of novel analysis to TND. The issue of follow-up and censoring has often been ignored in many studies

by using a binary indicator for COVID outcomes after vaccination, but a time-to-event model is needed to properly account for censoring and difference in follow-up time (33). To our knowledge, standard survival analysis methods, such as the renowned Cox regression model, have not been applied to evaluate VE with data from a TND study, but theoretical justification can be found in well-established case-control study literature (34). For such an analysis, the follow-up of test-negative controls may be right censored at a prespecified calendar time. More assumptions need to be made such that the comparison between the test-positive and test-negative groups reflects the VE in the target population. An example of such an assumption may be that the vaccination does not change the distribution of time to the symptom onset of test-negative illnesses.

With better data, better design, and better analysis with clearly stated assumptions, we can improve our understanding of key therapeutic and public-health questions around COVID-19 vaccines and risk factors.

REFERENCES

1. Gu T, Mack JA, Salvatore M, et al. Characteristics associated with racial/ethnic disparities in COVID-19 outcomes in an academic health care system. *JAMA Netw Open.* 2020; 3(10):e2025197.
2. Lipsitch M, Jha A, Simonsen L. Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies. *Int J Epidemiol.* 2016;45(6): 2060–2074.
3. 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):345–353.
4. Vandenbroucke JP, Pearce N. Test-negative designs: differences and commonalities with other case-control studies with "other patient" controls. *Epidemiology.* 2019;30(6): 838–844.
5. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine.* 2013; 31(17):2165–2168.
6. Lewnard JA, Patel MM, Jewell NP, et al. Theoretical framework for retrospective studies of the effectiveness of SARS-CoV-2 vaccines. *Epidemiology.* 2021;32(4): 508–517.
7. Lewnard JA, Cobey S. Immune history and influenza vaccine effectiveness. *Vaccine.* 2018;6(2):28.
8. Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun.* 2020;11(1):5479.
9. Dean NE, Hogan JW, Schnitzer ME. Covid-19 vaccine effectiveness and the test-negative design. *N Engl J Med.* 2021;385(15):1431–1433.
10. Dean NE, Halloran ME, Longini IM Jr. Temporal confounding in the test-negative design. *Am J Epidemiol.* 2020;189(11):1402–1407.
11. Stowe J, Tessier E, Zhao H, et al. Interactions between SARS-CoV-2 and influenza, and the impact of coinfection on disease severity: a test-negative design. *Int J Epidemiol.* 2021;50(4):1124–1133.
12. Lipsitch M. Challenges of vaccine effectiveness and waning studies. *Clin Infect Dis.* 2019;68(10):1631–1633.
13. Kahn R, Schrag SJ, Verani JR, et al. Identifying and alleviating bias due to differential depletion of susceptible people in postmarketing evaluations of COVID-19 vaccines. *Am J Epidemiol.* 2022;191(5):800–811.
14. Accorsi EK, Britton A, Fleming-Dutra KE, et al. Association between 3 doses of mRNA COVID-19 vaccine and symptomatic infection caused by the SARS-CoV-2 Omicron and Delta variants. *JAMA.* 2022;327(7):639–651.
15. Andrews N, Stowe J, Kirsebom F, et al. Covid-19 vaccine effectiveness against the Omicron (B.1.1.529) variant. *N Engl J Med.* 2022;386(16):1532–1546.
16. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Erratum: negative controls: a tool for detecting confounding and bias in observational studies (*Epidemiology* (2010) 21 (383–388)). *Epidemiology.* 2010;21(3):383–388.
17. Dagan N, Barda N, Kepten E, et al. BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *N Engl J Med.* 2021;384(15):1412–1423.
18. 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 Heal Am.* 2021;1:100025.
19. Vandenbroucke JP, Brickley EB, Vandenbroucke-Grauls CMJE, et al. A test-negative design with additional population controls can be used to rapidly study causes of the SARS-CoV-2 epidemic. *Epidemiology.* 2020;31(6):836–843.
20. Vandenbroucke JP, Brickley EB, Pearce N, et al. The evolving usefulness of the test-negative design in studying risk factors for COVID-19. *Epidemiology.* 2022;33(2):e7–e8.
21. El Adam S, Zou M, Kim S, et al. SARS-CoV-2 mRNA vaccine effectiveness in health care workers by dosing interval and time since vaccination: test-negative design, British Columbia, Canada. *Open ForumInfect Dis.* 2022; 9(5):ofac178.
22. Hitchings MDT, Ranzani OT, Lind ML, et al. Change in covid-19 risk over time following vaccination with CoronaVac: test negative case-control study. *BMJ.* 2022; 377:e070102.
23. Fogh K, Strange JE, Scharff BFSS, et al. Testing Denmark: a Danish nationwide surveillance study of COVID-19. *Microbiol Spectr.* 2021;9(3):e01330–e01321.

24. Allen WE, Altae-Tran H, Briggs J, et al. Population-scale longitudinal mapping of COVID-19 symptoms, behaviour and testing. *Nat Hum Behav.* 2020;4(9): 972–982.

25. Tleyjeh IM, Kashour T, Mandrekar J, et al. Overlooked shortcomings of observational studies of interventions in coronavirus disease 2019: an illustrated review for the clinician. *Open Forum Infect Dis.* 2021;8(8):ofab317.

26. Tenforde MW, Self WH, Gaglani M, et al. Effectiveness of mRNA vaccination in preventing COVID-19–associated invasive mechanical ventilation and death—United States, March 2021–January 2022. *Morb Mortal Wkly Rep.* 2022; 71(12):459–465.

27. McKeigue PM, Burgul R, Bishop J, et al. Association of cerebral venous thrombosis with recent COVID-19 vaccination: case-crossover study using ascertainment through neuroimaging in Scotland. *BMC Infect Dis.* 2021; 21(1):1275.

28. Thompson MG, Stenehjem E, Grannis S, et al. Effectiveness of Covid-19 vaccines in ambulatory and inpatient care settings. *N Engl J Med.* 2021;385(15):1355–1371.

29. Olson SM, Newhams MM, Halasa NB, et al. Effectiveness of BNT162b2 vaccine against critical Covid-19 in adolescents. *N Engl J Med.* 2022;386(8):713–723.

30. World Health Organization. Estimating COVID-19 vaccine effectiveness against severe acute respiratory infections (SARI) hospitalisations associated with laboratory-confirmed SARS-CoV-2: an evaluation using the test-negative design: guidance document. https://covid19.recmap.org/guideline/p_l_dudasj_mcmaster_ca_0_abcf4032-54f4-4788-bd04-ee5c38ae8328. 2021. Accessed December 2, 2022.

31. Chung H, He S, Nasreen S, et al. Effectiveness of BNT162b2 and mRNA-1273 covid-19 vaccines against symptomatic SARS-CoV-2 infection and severe covid-19 outcomes in Ontario, Canada: test negative design study. *BMJ.* 2021; 374:n1943.

32. Andrews N, Stowe J, Kirsebom F, et al. Effectiveness of COVID-19 booster vaccines against COVID-19-related symptoms, hospitalization and death in England. *Nat Med.* 2022;28(4):831–837.

33. Lin D-Y, Gu Y, Wheeler B, et al. Effectiveness of Covid-19 vaccines over a 9-month period in North Carolina. *N Engl J Med.* 2022;386(10):933–941.

34. Prentice RL, Breslow NE. Retrospective studies and failure time models. *Biometrika.* 1978;65(1):153–158.

35. Centers for Dease Control and Prevention. Covid Data Tracker. https://covid.cdc.gov/covid-data-tracker/#trends_dailycases. Accessed December 2, 2022.