1 **When emulating a trial, do as the trialists do: Missteps in estimating relative effectiveness**

2 **of a SARS-CoV-2 vaccine booster dose**

3

4 Kathleen E. Wirth, ScD[1]

5 Jessie K. Edwards, PhD[2]

6 Lydia Feinstein, PhD[1]

7 Alexander Breskin, PhD[1]

8

9

10 [1]NoviSci, a division of Target RWE Health Evidence Solutions, Durham, NC, USA

11 [2]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC,

12 USA.

13

14

15 Corresponding author:

16 Kathleen E Wirth, ScD

17 Target RWE Health Evidence Solutions

18 Durham, NC UNITED STATES

19 Email: kwirth@hsph.harvard.edu

20

1    Dear Editor,

2    We read with interest the recent study by Butt A.A. et al [1] "Relative vaccine effectiveness of a
3    SARS-CoV-2 mRNA vaccine booster dose against the omicron variant." This observational
4    study, conducted in the Department of Veteran Affairs Healthcare System, estimated the relative
5    effectiveness of a booster dose for preventing confirmed CoV-2 infection (19%), hospitalization
6    (52%), and intensive care unit admission or death (83%). The authors state that they emulated a
7    target trial, a framework that structures the research question and mitigates common sources of
8    bias [2–5]. Unfortunately, the authors did not define or adhere to the target trial they sought to
9    emulate, leading to several study design and analysis concerns.

10   First, the authors assigned treatment status before determining eligibility, reversing the expected
11   order of operations. Among vaccinated individuals, the authors identified those who received a
12   booster between 09/22/21 and 12/25/2021, and then matched booster recipients with individuals
13   who had not yet received a booster. If either member of the pair became infected with SARS-
14   CoV-2 prior to the start of follow-up on 01/01/2022, both individuals were excluded. This
15   approach is problematic because it 1) requires individuals to remain free of infection for up to
16   three months between becoming eligible for and starting follow-up, and 2) introduces differential
17   selection pressures by treatment status. For example, among unboosted people, the successful
18   avoidance of SARS-CoV-2 infection during the months of peak omicron activity in the United
19   States suggests the presence of an unmeasured protective factor (perhaps strict masking and/or
20   social distancing). Such confounding factors would likely lead to an underestimation of booster
21   effectiveness. Similarly, if effectiveness wanes over time, delaying follow-up for up to three
22   months unnecessarily discards the most valuable "boosted" person-time, further underestimating
23   booster effectiveness. In an appropriately conducted trial, there is no waiting period.

24   The authors' findings are further compromised by missteps in their analytic approach. First, the
25   authors censored controls who received a booster during follow-up without implementing well
26   defined statistical approaches (e.g., weighting) to mitigate selection bias likely induced by
27   informative censoring [6,7]. This is problematic because controls who received a booster during
28   follow-up may systematically differ from those who remained unboosted with respect to their
29   risk for infection and severe disease. Second, the authors' analytic approach did not fully adjust
30   for the bias they sought to address via matching. Although the authors carefully matched case
31   and control subjects at baseline, they did not subsequently conduct a matched analysis using
32   pair-stratified regression. Such an analysis controls for both the main effects of the matching
33   factors, but also any interactions between factors. This analytic misstep is particularly concerning
34   in the current context as sex, age, and underlying medical conditions independently and jointly
35   impact SARS-CoV-2-related outcomes [8,9].

36   Target trial emulation provides a framework for making causal inference from observational
37   data. Unfortunately, the authors did not explicitly define their target trial and did not adhere to
38   the principles of the overall framework. These lapses open the door for substantial bias and,
39   unfortunately, limit meaningful interpretation of the study's findings.

40

1 **NOTES**

2 **Financial Support**: None

3

4 **Conflicts of interest**: KEW, LF, and AB are employees of Target RWE. LF and AB are
5 shareholders of Target RWE. JE reports grants or contracts from NIH (paid to institution),
6 consulting fees from Plos Medicine (paid to author), support for attending meetings and/or travel
7 from Society for Epidemiologic Research (paid to author), and is Leadership or fiduciary role in
8 other board, society, committee or advocacy group for Society of Epidemiologic Research
9 (unpaid participation)

10

**References:**

1. Butt A.A., Talisa V.B., Shaikh O.S., Omer S.B., Mayr F.B. Relative Vaccine Effectiveness of a SARS-CoV-2 mRNA Vaccine Booster Dose Against the Omicron Variant. *Clinical Infectious Diseases*, 2022; ciac328, https://doi.org/10.1093/cid/ciac328.

2. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016; 183:758–764.

3. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016; 79:70–75.

4. García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol* 2017; 32:495–500.

5. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019; 25:1601–1606.

6. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics.* 2000; 56:779–788.

7. Willems SJW, Schat A, van Noorden MS, Fiocco M. Correcting for dependent censoring in routine outcome monitoring data by applying the inverse probability censoring weighted estimator. *Stat Methods Med Res*. 2018; 27:323–335.

8. Kompaniyets L, Pennington AF, Goodman AB, Rosenblum HG, Belay B, Ko JY, et al. Underlying Medical Conditions and Severe Illness Among 540,667 Adults Hospitalized With COVID-19, March 2020–March 2021. *Prev Chronic Dis*. 2021; 18:210123.

9. Pennington AF, Kompaniyets L, Summers AD, Danielson ML, Goodman AB, Chevinsky JR, et al. Risk of Clinical Severity by Age and Race/Ethnicity Among Adults Hospitalized for COVID-19—United States, March–September 2020. *Open Forum Infect Dis.* 2020; 8(2):ofaa638.