Rejoinder: The Future of Outcome-Wide Studies

Tyler J. VanderWeele, Maya B. Mathur and Ying Chen

We thank Daniel (2020), Vansteelandt and Dukes (2020) and Ertefaie and Johnson (2020) for their thoughtful, insightful and enlightening commentaries on our paper (VanderWeele, Mathur and Chen, 2020). We have learned a great deal from their comments, discussion and proposals. The outcome-wide approach is still in its infancy and, as pointed out by the commentators, there are certainly numerous ways to refine and extend what we had proposed as a basic template. The analytic approaches even to estimate causal effects of a single time-fixed exposure on a single subsequent outcome have increased dramatically over the past decades. The range of considerations and decisions that arise when considering multiple outcomes are yet more vast, and thus, over time, there may likewise be an array of principled analytic options for outcome-wide studies, too. We suspect that many of the seeds for that potentially vast array of options are likely to be found in the commentaries of Daniel (2020), Vansteelandt and Dukes (2020) and Ertefaie and Johnson (2020). We will respond to their various remarks by considering how they give rise to important cautions, important extensions and important alternatives to the practice of outcome-wide studies.

IMPORTANT CAUTIONS

The commentators raise a range of important points and caveats to the implementation and interpretation of the outcome-wide analytic approach that we proposed, some of which indeed may not have received due attention in our paper. Daniel (2020) rightly points out that when, with covariate data that is contemporaneous with the exposure, it is unclear whether a particular covariate is a confounder or a mediator, the approach of considering analyses both with and without the covariate will not necessarily bound the causal effect. The two analyses can

Tyler J. VanderWeele is John L. Loeb and Frances Lehman Loeb Professor of Epidemiology, Harvard T.H. Chan School of Public Health, and Institute for Quantitative Social Science, Harvard University, Boston, Massachusetts 02115, United States (e-mail: tvanderw@hsph.harvard.edu). Maya B. Mathur is Assistant Professor, Quantitative Sciences Unit, Stanford University, Palo Alto, California 94304, United States. Ying Chen is Research Associate, Institute for Quantitative Social Science, Harvard University, Cambridge, Massachusetts 02138, United States.

be biased in the same direction when the covariate's effect on the exposure is of the opposite direction of the exposure's effect on the covariate. While we do still think considering both analyses with and without the covariate is valuable, we certainly acknowledge that concordance of these two analyses does not necessarily indicate a clear conclusion. Discordance should raise cause for concern; but even with concordance one should be cautious in interpretation. When the temporal structure of the data is such that covariate levels prior to the exposure cannot be adjusted for, and it is thus unclear whether a covariate is a confounder or a mediator, we think it will often be difficult to draw causal conclusions (VanderWeele, 2015).

Daniel (2020) also rightly points out the potential importance, if one is using multiple imputation to handle missing data, of including all outcomes simultaneously in imputation models. In our early work implementing the outcome-wide approach, we had indeed neglected this (Chen and VanderWeele, 2018; Chen et al., 2019a, 2019b), but in all of our more recent empirical outcome-wide analyses (Chen et al., 2019c, Chen, Kubzansky and VanderWeele, 2019, Long et al., 2020, Kim et al., 2020), this is indeed how we have proceeded. The analytic approach to these outcome-wide studies is certainly still evolving, and there will almost certainly be other refinements to it, like this one, a point to which we will also return below.

Daniel (2020) raises further important concerns about potential positivity violations. The analysis, regardless of whether using propensity scores, or regression models, or doubly-robust methods, requires that the groups with and without exposure have overlap in the covariates. If certain exposure or treatment decisions are made deterministically then this can be violated. While this is indeed a well-known fact within causal inference, there was arguably not sufficient emphasis of it in our paper, and as Daniel (2020) rightly notes, the problem may be compounded by the large number of covariates for which adjustment might be made in an outcome-wide approach. Checking adequate covariate overlap can often be facilitated by estimating propensity scores, an alternative analytic approach for outcome-wide studies discussed in our paper, and also advocated for in the outcome-wide context by Vansteelandt and Dukes (2020), and which we will also consider further below.