American Journal of Epidemiology

Vol. 188, No. 7
DOI: 10.1093/aje/kwy281
Advance Access publication:
December 24, 2018

Point-Counterpoint

Counterpoint: Mediation Formulas With Binary Mediators and Outcomes and the "Rare Outcome Assumption"

Tyler J. VanderWeele*, Linda Valeri, and Cande V. Ananth

* Correspondence to Dr. Tyler J. VanderWeele, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115 (e-mail: tvanderw@hsph.harvard.edu).

Initially submitted December 4, 2018; accepted for publication December 11, 2018.

In their accompanying article, Samoilenko and Lefebvre (*Am J Epidemiol*. 2019;188(7):1203–1205) correctly note 2 typographical errors in the formulas presented in a 2011 paper on placental abruption by Ananth and VanderWeele (*Am J Epidemiol*. 2011;174(1):99–108). Fortunately, to the best of our knowledge, researchers are using our methods papers and Dr. VanderWeele's 2015 book on mediation analysis (*Explanation in Causal Inference*; Oxford University Press, New York, New York), rather than the paper on placental abruption, to carry out their direct and indirect effect analyses; and in our methods papers and the book, the formulas are correct. The formulas discussed by Samoilenko and Lefebvre and in our work make reference to a "rare outcome assumption." In evaluating this assumption, it is important to note that the outcome is to be relatively rare *across all strata defined by the exposure and the mediator*—a point that is often neglected.

direct and indirect effects; mediation; odds ratios; rare outcomes

***Editor's note**: A counterpoint to this article appears on page 1201.*

We thank Samoilenko and Lefebvre (1) for noticing 2 typographical errors in the 2011 paper by Ananth and VanderWeele (2) entitled, "Placental Abruption and Perinatal Mortality With Preterm Delivery as a Mediator: Disentangling Direct and Indirect Effects." They are correct that there were typographical errors, both with the coefficient in the denominator of the risk ratio for the natural direct effect (RR_{NDE}) and with the use of "log" rather than "logit" in the model for the mediator. We regret not having noticed these errors earlier. Fortunately, to the best of our knowledge, researchers have relied not on this empirical application paper for the formulas for direct and indirect effects but upon methodological papers, such as the *Psychological Methods* paper of Valeri and VanderWeele (3) and the book by VanderWeele (4), for their analyses. In both the paper (3) and the book (4), the formulas are correct, and the model for the mediator is a logit model. In the 80 citations (as reported in Google Scholar (scholar.google.com; Google Inc., Mountain View, California); assessed on December 4, 2018) of the Ananth and VanderWeele article (2), virtually all concern the substantive points made in the paper, rather than the formulas therein. With 547 citations of the paper by Valeri and VanderWeele (3) and 444 citations of the book by VanderWeele (4), many of these *are* applications of the relevant formulas; and again, in those methodological references, the formulas are correct.

We regret that Samoilenko and Lefebvre (1) had trouble finding the supplementary material for the paper by Valeri and VanderWeele (3). The material is indeed difficult to find on the *Psychological Methods* website, but it is available there (5). We had previously raised concerns about the lack of its visibility to the journal's staff, but sadly they did not seek to fix the issue. The supplementary materials have also been available on Dr. VanderWeele's website (6) for several years.

Like Samoilenko and Lefebvre (1), we had also pointed out previously (4) that with the logit model, the mediator need not be rare for the formulas to apply; only the outcome needs to be rare.

We also appreciate the work that Samoilenko et al. (7) have done elsewhere on examining different approximations for direct and indirect effect risk ratios. Not all approximations are equal, and some may work better in certain scenarios than others (8). Note also that in many settings, including the context of direct and indirect effects, the "rare outcome assumption" should hold across all of the relevant strata of the exposure (and mediator) variable(s) for the odds ratio to necessarily be a good approximation of the risk ratio (9). Regrettably, neither we nor Samoilenko and colleagues, nor many other authors, have been as careful in the precise statement of this assumption as might be hoped. Notably, in scenario 2 in the paper by Samoilenko et al. (7), the probability of the outcome when the exposure and

mediator are present is 70% or more—certainly not rare!—and it is perhaps unsurprising, then, that the approximations do not work as well as might be desired. Greater care should certainly be taken in the articulation of this assumption, and we will endeavor to do so in the future. The rare outcome assumption, across strata of the exposure and mediator, is still quite plausible in our 2011 application (2), where the outcome was perinatal mortality; but it is certainly not plausible in the application of Samoilenko et al. (7), where the outcome was low birth weight, for which there was a probability of 70.4% when both the exposure (abruption) and the mediator (preterm delivery) were present. However, the fact that the approximations they proposed do somewhat better, even in this scenario, is indeed of interest.

ACKNOWLEDGMENTS

Author affiliations: Departments of Epidemiology and Biostatistics, T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts (Tyler J. VanderWeele); Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York (Linda Valeri); Division of Epidemiology and Biostatistics, Department of Obstetrics, Gynecology, and Reproductive Sciences, Robert Wood Johnson Medical School, Rutgers University, New Brunswick, New Jersey (Cande V. Ananth); and Environmental and Occupational Health Sciences Institute, Robert Wood Johnson Medical School, Rutgers University, Piscataway, New Jersey (Cande V. Ananth).

This work was funded by the National Institutes of Health (grant R01 CA222147).

Conflict of interest: none declared.

REFERENCES

1. Samoilenko M, Lefebvre G. Point: risk ratio equations for natural direct and indirect effects in causal mediation analysis of a binary mediator and a binary outcome—a fresh look at the formulas. *Am J Epidemiol*. 2019;188(7):1201–1203.
2. Ananth CV, VanderWeele TJ. Placental abruption and perinatal mortality with preterm delivery as a mediator: disentangling direct and indirect effects. *Am J Epidemiol*. 2011;174(1):99–108.
3. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*. 2013;18(2):137–150.
4. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York, NY: Oxford University Press; 2015.
5. American Psychological Association. Supplemental material for mediation analysis allowing for exposure–mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. http://supp.apa.org/psycarticles/supplemental/a0031034/a0031034_supp.html. Accessed December 4, 2018.
6. Harvard T.H. Chan School of Public Health. Tyler VanderWeele. Tools and tutorials. www.hsph.harvard.edu/tyler-vanderweele/tools-and-tutorials/. Accessed December 4, 2018.
7. Samoilenko M, Blais L, Lefebvre G. Comparing logistic and log-binomial models for causal mediation analyses of binary mediators and rare binary outcomes: evidence to support cross-checking of mediation results in practice. *Obs Stud*. 2018;4:193–216.
8. VanderWeele TJ. On a square-root transformation of the odds ratio for a common outcome [letter]. *Epidemiology*. 2017;28(6):e58–e60.
9. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.