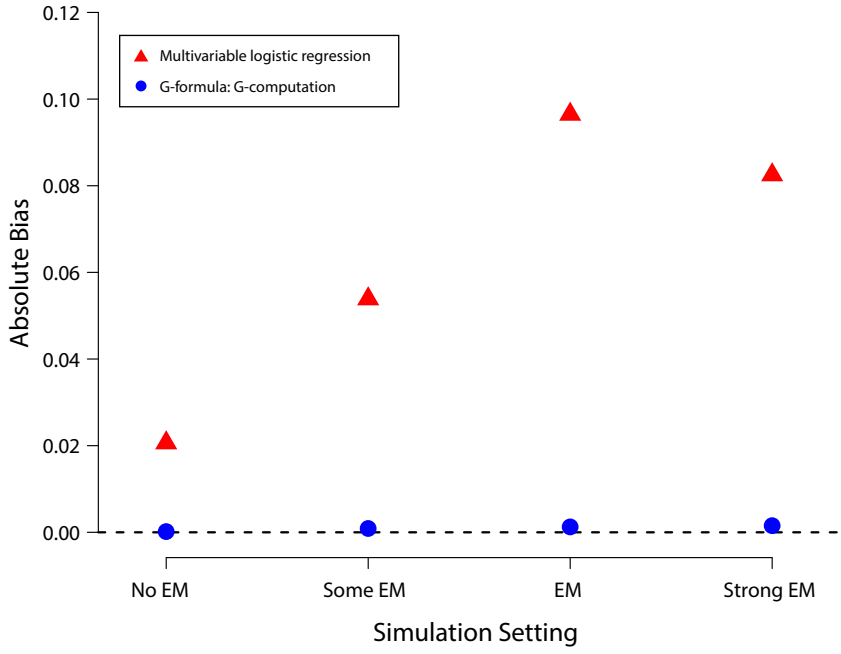
EFFECT MODIFICATION AND COLLAPSIBILITY IN EVALUATIONS OF PUBLIC HEALTH INTERVENTIONS

The Evaluating Public Health Interventions *AJPH* series offers excellent practical guidance to public health researchers. The eighth part of the series provides a valuable introduction to effect estimations of time-invariant public health interventions.[1] In their commentary Spiegelman and Zhou suggest that, in terms of bias and efficiency, there is no advantage to using modern causal inference methods over classical multivariable modeling.[1] However, this statement is not always true. Most important, both effect modification and collapsibility are critical concepts when assessing the validity of using regression for causal effect estimation.

Suppose that one is interested in the effect of combined radiotherapy and chemotherapy versus chemotherapy only on one-year mortality among patients diagnosed with colorectal cancer. A clinician may ask: how different would the risk of death have been had everyone received dual therapy as compared with if everyone had experienced monotherapy? The causal marginal odds ratio (MOR) offers an answer to this question. Each individual has a pair of potential outcomes: the outcome he or she would have experienced had he or she been exposed to dual treatment (A = 1), denoted Y(1), and the outcome had he or she been unexposed, Y(0). The MOR is defined as $[P(Y(1) = 1)/(1 - P(Y(1) = 1))]/[P(Y(0) = 1)/(1 - P(Y(0) = 1))]$.

A common approach would be to use logistic regression to model the odds of mortality given the intervention and adjust for confounders (W) such as clinical stage and comorbidities. Note that this regression will provide an estimate of the conditional odds ratio (COR), which is $[P(Y = 1 \mid A = 1,W) / (1 - P(Y = 1 \mid A = 1,W))] / [P(Y = 1 \mid A = 0,W) / (1 - P(Y = 1 \mid A = 0,W))]$. The MOR and COR are typically not identical. First, if there is effect modification (e.g., if the effect of dual therapy is different between patients with no comorbidities and those who have hypertension), logistic regression including an interaction term will not provide a marginal effect estimate but only the conditional effect of the interaction term between dual therapy and hypertension. Second, the odds ratio is noncollapsible, which means that the MOR is not necessarily equal to the stratum-specific odds ratio (i.e., the COR). This holds even when a covariate is related to the outcome but not the intervention and is thus not a confounder.[2,3]

Figure 1 shows the results of a simulation based on our cancer example. The figure compares bias with respect to the MOR for logistic regression versus the G-formula[4,5] (a causal inference method) for different levels of effect modification. The details of the set-up can be found online.[6] One can see the bias of logistic regression, which is more pronounced



Note. EM = effect modification. The graph represents a comparison of the mortality risk difference one year after cancer diagnosis among patients on dual therapy (radiotherapy and chemotherapy) versus patients on monotherapy (chemotherapy only). Known confounders are age, socioeconomic status, comorbidities, and clinical stage. The absolute bias with respect to the marginal causal odds ratio is reported on the basis of a sample size of 5000 and 10 000 simulation runs (see https://github.com/migariane/hetmor).

FIGURE 1—Results of the Cancer Simulation

Letters to the editor referring to a recent *AJPH* article are encouraged up to 3 months after the article's appearance. By submitting a letter to the editor, the author gives permission for its publication in *AJPH*. Letters should not duplicate material being published or submitted elsewhere. The editors reserve the right to edit and abridge letters and to publish responses.

Text is limited to 400 words and 7 references. Submit online at www.editorialmanager.com/ajph. Queries should be addressed to the Editor-in-Chief, Alfredo Morabia, MD, PhD, at editorajph@qc.cuny.edu.

under effect modification but persists—as a result of noncollapsibility—even under no effect modification. AJPH

Miguel Angel Luque-Fernandez, PhD, MA, MPH, MSc
Daniel Redondo-Sanchez
Michael Schomaker, PhD

ABOUT THE AUTHORS

Miguel Angel Luque-Fernandez is with the Department of Noncommunicable Disease and Cancer Epidemiology, University of Granada, Granada, Spain, and the Department of Noncommunicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, England. Daniel Redondo-Sanchez is with Department of Noncommunicable Disease and Cancer Epidemiology, University of Granada, Granada, Spain. Michael Schomaker is with the Center for Infectious Disease Epidemiology and Research, School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa.

Correspondence should be sent to Miguel Angel Luque-Fernandez, PhD, MA, MPH, MSc, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom (e-mail: miguel-angel.luque@lshtm.ac.uk). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This letter was accepted December 3, 2018.
doi: 10.2105/AJPH.2018.304916

CONTRIBUTORS

M. A. Luque-Fernandez wrote the letter. M. A. Luque-Fernandez and D. Redondo-Sanchez set up the repository. M. A. Luque-Fernandez and M. Schomaker carried out the simulations. All of the authors interpreted the data and drafted and revised the letter.

ACKNOWLEDGMENTS

Miguel Angel Luque-Fernandez is supported by the Spanish National Institutes of Health (Carlos III Miguel Servet I Investigator Award CP17/00206).

We thank Mary-Ann Davies, Leigh Johnson, Morna Cornell, and Landon Myer for their feedback on this letter.

CONFLICTS OF INTEREST

No conflicts of interest.

REFERENCES

1. Spiegelman D, Zhou X. Evaluating public health interventions: 8. Causal inference for time-invariant interventions. *Am J Public Health*. 2018;108(9)1187–1190.

2. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci*. 1999;14(1): 29–46.

3. Sjölander A, Dahlqwist E, Zetterqvist J. A note on the noncollapsibility of rate differences and rate ratios. *Epidemiology*. 2016;27(3):356–359.

4. Luque-Fernandez MA, Schomaker M, Rachet B, Schnitzer ME. Targeted maximum likelihood estimation for a binary treatment: a tutorial. *Stat Med*. 2018;37(16): 2530–2546.

5. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol*. 2011; 173(7):731–738.

6. Luque-Fernandez MA, Redondo-Sanchez D, Schomaker M. Effect modification and collapsibility when estimating the effect of interventions: a Monte Carlo simulation comparing classical multivariable regression adjustment versus the G-formula based on a cancer epidemiology illustration. Available at: https://github.com/migariane/hetmor. Accessed December 18, 2018.

SPIEGELMAN AND ZHOU RESPOND

In their letter, Luque-Fernandez et al. considered the impact of effect measure modification and collapsibility on the estimation of causal effects of time-invariant public health interventions under time-invariant exposures, using a standard multivariate logistic regression model. The primary message of our column "Evaluating Public Health Interventions" was that, contrary to what may be an increasingly widespread misperception, it is nearly always the case that standard multivariate modeling approaches will provide valid causal estimates of time-invariant interventions. Because public health interventions and programs are usually time-invariant, this message is particularly salient for readers of this journal.

Our detailed responses to their letter follow.

1. Can multivariate models be used for causal inference about time-invariant public health interventions in the presence of effect modification? First, in practice, effect modification is rare on the multiplicative scale—that is, in logistic regression, log-linear regression, and Cox regression, as previously discussed by Spiegelman, VanderWeele, and others in this journal.[1,2] When effect modification is present, the effect modifiers need to be included in the model, and it has been recommended for at least the past 40 years to report effects by each level of the modifiers. In the rare cases when there is effect modification that is ignored, the model is incorrect and, of course, bias will result, as is always the case when misspecified models are to attempt to estimate causal parameters.

In contrast to this time-tested approach for considering effect modification, Luque-Fernandez et al. studied by simulation the extent to which the causal odds ratio is biased when effect modification is present but ignored. We likely all know that in this circumstance, the estimated effect measure is not interpretable. I was teaching this point in my intermediate epidemiology methods course (EPI202) at the Harvard School of Public Health as early as 1993, and I believe I was taught this in the same course by Ken Rothman five years previously. It is disappointing when such fundamental principles of epidemiological research seem to be forgotten and a letter like this appears.

Luque-Fernandez et al. simulated examples of the value of the (uninterpretable) estimated odds ratio under mild to strong effect modification that is ignored. To the extent that these simulations could have been of any interest, they used a "straw man" approach, where the effect modification considered produced stratum-specific relative risks as extreme as 0.08 and 0.13, in the moderate and strong cases, respectively. We would again like to emphasize that multiplicative effect modification is rarely seen in practice, and in the rare cases that we are aware of when it has been reliably reported, the magnitude of the effect has been nowhere close to the extreme cases considered by Luque-Fernandez et al.

2. Much has been made of the noncollapsibility property of the odds ratio. Far too much, in our view. Although theoretically there is no question that the odds ratio is indeed noncollapsible, again, it is rarely the case that this matters in wide practice.[3] As shown in Nevo et al.,[4] when the outcome is not rare, a strong relative risk is required for noncollapsibility bias to be severe. When the outcome is rare, the conditions under which noncollapsibility bias appreciably appears need to be even more extreme. In the first simulated example of Luque-Fernandez et al., where there was no effect modification, restricting the question at hand to only that of the impact of collapsibility, the value of the odds ratio was extreme at about 0.075. With an absolute bias of 0.02, the standard estimated odds ratio would be overestimated at 0.055. It is unlikely that any substantive interpretation of such an extreme effect of an intervention would change as a result of this bias. In any case, such a strong relative risk is rarely observed in practice. The "straw man" argument has been used here to propel undue concerns about the collapsibility issue well beyond their usefulness in real-life public health evaluations.