Interactions in Epidemiology: Relevance, Identification, and Estimation

Sander Greenland

Sufficient-cause models represent biologic interactions (including synergism) in the form of coparticipation of factors in a sufficient cause[1–3] (causal coaction). In this issue of the journal, VanderWeele[1] provides new results relating sufficient-cause models to regression models. These results reinforce the point that concepts of biologic interaction do not in general correspond to the concept of statistical interaction, because the latter is only the need for a product term in a statistical model.[4,5] They further underscore that, without strict assumptions, sufficient-cause interactions do not correspond to interdependence of effects in potential-outcomes (counterfactual) models.

I will examine the practical relevance and limitations of using epidemiologic regressions to study interactions, however the latter are defined.

Relevance of Interactions

A practical issue concerning interactions is their relevance to public health goals of finding cost-effective intervention strategies for disease reduction. Part of public health involves identifying population subgroups that would benefit most from a given intervention. These are not always high-risk subgroups—it is possible that a lower-risk group would obtain the greatest absolute risk reduction from the proposed intervention.

In the absence of bias, departures from risk additivity imply that some subgroups would obtain a greater absolute risk reduction from the intervention than others would. Thus earlier authors noted that if costs were measured in terms of case-load per unit population (average risk), the relevant null model for detecting special groups would be additive on the risk scale.[6,7] Risk nonadditivity (or "public health interaction") implies the existence of such groups.[5,6] It also implies existence of interdependent effects in individuals, also known as "interactions" in basic potential-outcome (counterfactual) causal models; that is, risk nonadditivity implies there are individuals such that the presence or direction of effect of one factor depends on the other factor.[5,8]

Risk nonadditivity is equivalent to heterogeneity or modification of risk differences for one variable across levels of another.[5] For factors that can act only causally, excess risk above additivity (superadditivity) among those exposed to both factors signals the presence of individuals who get the disease only when exposed to both factors ("synergism" in the potential-outcomes sense).[4,5,8] Identifying such individuals is valuable for preventive medicine and public health, because for them the outcome can be prevented by removing either of the factors.

Nonidentifiability of Biologic Interactions

Epidemiologic studies observe population distributions of risks, which submerge details of individual experiences. One consequence of this submersion is that many different biologic models will lead to the same population model.[9,10] In particular, the implications of biologic models for interaction are unidirectional; for

From the Departments of Epidemiology and Statistics, University of California, Los Angeles, Los Angeles, CA.
Correspondence: Sander Greenland, Departments of Epidemiology and Statistics, University of California, Los Angeles, CA 90095-1772. E-mail: lesdomes@ucla.edu.

example, risk additivity could hold because no interdependence of effects exists, or instead may hold because of cancellations of the population effects of different individual interactions.[5,8]

Perfect cancellation of interactions is a higher-order analog of "unfaithfulness" in causal diagrams,[11,12] where 2 variables may be unassociated despite having causal connections. For example, a treatment with effects could appear independent of the outcome if it causes and prevents equal numbers of cases. Although perfect cancellation is often implausible, any degree of cancellation will make it more difficult to detect effects or their interdependencies.

Even if we could observe each individual's response under each exposure pattern, we still could not determine from those observations which biologic mechanisms or sufficient-cause types were operative[8,13–15]; hence, individual mechanisms are nonidentifiable even from a perfect factorial or crossover trial.[13,15] For example, even with no sufficient-cause synergism, variation in background characteristics could nonetheless produce superadditivity.[14] More generally, estimation of degrees of interaction will require further assumptions (eg, monotonicity) that cannot be tested with epidemiologic data, even if the latter are perfect, although valid tests for interactions can be constructed from weaker conditions.[1,5,8,13–15]

Statistical Modeling Issues

Epidemiologic theories of interactions[1,4,5,8,9,14–17] derive population causal models from models for individual etiology (eg, risk additivity is derived from the absence of interdependent effects). Study of these models and forms does not require that one use the derived population model as the statistical analysis model. One can estimate risks or rates by using any statistical model, then use these estimates to test or estimate the parameters in the causal model.[1,18] For example, one can use risk estimates from a logistic statistical model for inference on parameters in an additive-risk causal model.[1]

Nonetheless, valid inference from a statistical model requires that the model have enough higher-order terms (such as products and nonlinear trend terms) to capture the actual risk or rate pattern. Unfortunately, most statistical practice mandates excluding terms that fail a significance test or similar criterion for inclusion. Because epidemiologic data typically have limited power to detect higher-order terms such as products,[19–21] the consequence of such criteria will be frequent exclusion of these terms. This exclusion in turn biases risk estimates toward the "main-effects only" form of the fitted statistical model.

As examples of prime concern, excluding terms from logistic, log-linear (Poisson regression), or proportional-hazards (Cox) models will result in risk estimates biased toward multiplicative effects and exponential dose-response for odds or rates. Such model forms are highly "risk nonadditive" in typical settings, so parsimony with terms induces a bias toward risk nonadditivity. Potential consequences of this statistical bias include overestimation of nonadditivity (leading to mistakes in targeting subgroups for intervention), overestimation of effect interdependence, and biased inferences about sufficient-cause interactions.

Switching to an additive-risk model results in analogous problems. Excluding higher-order terms from additive-risk models produces bias toward risk additivity and linear dose-response. This bias can result in underestimation of risk nonadditivity (again leading to mistakes in targeting subgroups) and underestimation of effect interdependence. Furthermore, when using stratified or conditional models (eg, to control matching factors), additive relative-risk models will no longer be valid substitutes for additive-risk models.[22] Consequently, to study nonadditivity, most case-control data will have to be modeled using background information about stratum-specific risks or rates to reconstruct population risk patterns.[23,24]

A straightforward way to avoid unacceptable bias from model simplification is to reorient statistical modeling away from parsimony and toward smoothing to identify patterns amid noise.[25,26] This reorientation allows one to stay within convenient multiplicative-model families such as the logistic, albeit with more complex model terms than are common in epidemiology. When many adjustment variables must be included, this goal leads to fitting methods that can use a very large set of terms, such as shrinkage[27] and machine-learning algorithms such as boosting.[28,29] The same methods can also be used to improve estimation of propensity scores for inverse-probability weighting in doubly robust estimation.[30] Accurate statistical inferences from these methods may require computer-intensive techniques such as bootstrapping (which involves more than just data resampling[31]).

Validity Issues

VanderWeele[1] notes that his results, like others, assume the absence of uncontrolled confounding. In practice we must also assume absence of uncontrolled selection bias and measurement error (including misclassification). Uncontrolled biases can spuriously mask, reverse, or generate heterogeneity (such as nonadditivity), and measurement error can cause any of these problems even if it is nondifferential.[21,32]

Unfortunately, few epidemiologic studies can control all sources of bias. For this reason, analyses that attempt to account honestly for uncertainties about biases can yield interval estimates that are far wider than those from conventional statistical models.[33,34] In large studies and meta-analyses, uncertainty due to random error can be far less than bias uncertainty, because random error in estimates shrinks with sample size but biases do not.[35] Because random error in interaction assessment will be larger,[19–21] there will be even larger total uncertainty than seen for average effects.

Terminology: Impact Beyond Words

In statistics, an "interaction" in a regression model is nothing more than a product term. Such terms can be made to disappear or even change sign simply by transforming the scale on which the outcome is measured (eg, by taking its logarithm). This terminology is unfortunate because the need for product terms or lack thereof does not correspond to the presence or absence of biologic interactions except in special cases,[1,8,15,16,21,36] which have been elaborated by Vander-Weele.[1,15] Thus, because there is no biologic rationale for such usage, calling product terms "interactions" is misleading and should be abandoned.

Sadly, one still sees confused reports claiming "interaction" is or is not present based on tests for departures from multiplicative models (eg, tests of product terms in logistic models), as if such results have biologic meaning in the absence of a causal model for interactions. Much of the literature on "gene-environment interactions" revolves around this confusion. Journal editors and reviewers could help abate this problem by requesting that authors of such analyses replace the word "interaction" by more precise descriptors such as "nonmultiplicativity" or "product term."

Conclusion

When (as is usually the case) adjustment for multiple covariates is required, valid inferences about risk nonadditivity and biologic interactions will require unconventional modeling approaches that replace parsimony by predictive accuracy as the statistical goal. Even if these approaches are deployed, I believe that (due to limited power and validity) only in exceptional circumstances will epidemiology be able to provide reliable inferences about biologic interactions, however those are defined. The chief exceptions will be situations in which the study factors have effects so large as to be undeniable, as with asbestos and smoking.[1]

In particular, VanderWeele[1] shows that, without further assumptions, the causal risk ratio for one causal factor must increase over 2-fold across the other causal factor to imply sufficient-cause interactions. Inferring such a condition requires a very large risk-ratio estimate for the jointly exposed from large amounts of excellent data. Even with huge effects and perfect data, however, identification of specific mechanisms of interaction will depend on biologic assumptions that are untestable (nonidentifiable) with epidemiologic observations alone.

REFERENCES

1. VanderWeele TJ. Sufficient cause interactions and statistical interactions. *Epidemiology*. 2009;20:6–13.
2. Rothman KJ, Greenland S, Poole C, Lash TL. Causation and causal inference. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008:51–70.
3. Rothman KJ. Causes. *Am J Epidemiol*. 1976;104:587–592.
4. Darroch J. Biologic synergism and parallelism. *Am J Epidemiol*. 1997;145:661–668.
5. Greenland S, Lash TL, Rothman KJ. Concepts of interaction. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008:71–83.
6. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol*. 1980;112:467–470.
7. Saracci R. Interaction and synergism. *Am J Epidemiol*. 1980;112:465–466.
8. Greenland S, Poole C. Invariants and noninvariants in the concept of interdependent effects. *Scand J Work Environ Health*. 1988;14:125–129.
9. Siemiatycki J, Thomas DC. Biological models and statistical interactions: an example from multistage carcinogenesis. *Int J Epidemiol*. 1981;10:383–387.
10. Thompson WD. Effect modification and the limits of biological inference from epidemiologic data. *J Clin Epidemiol*. 1991;44:221–232.
11. Greenland S, Pearl J. Causal diagrams. In: Boslaugh S, ed. *Encyclopedia of Epidemiology*, Vol. 1. Thousand Oaks, CA: Sage; 2007:149–156.
12. Glymour MM, Greenland S. Causal diagrams. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008:183–209.
13. Greenland S, Brumback B. An overview of relations among causal modelling methods. *Int J Epidemiol*. 2002;31:1030–1037.
14. VanderWeele TJ, Robins JM. The identification of synergism in the sufficient-component-cause framework. *Epidemiology*. 2007;18:329–339.
15. VanderWeele TJ, Robins JM. Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika*. 2008;95:49–61.
16. Koopman JS. Interaction between discrete causes. *Am J Epidemiol*. 1981;113:716–724.
17. Weinberg CR. Applicability of the simple independent action model to epidemiologic studies involving two factors and a dichotomous outcome. *Am J Epidemiol*. 1986;123:162–173.
18. Greenland S. Limitations of the logistic analysis of epidemiologic data. *Am J Epidemiol*. 1979;110:693–698.
19. Greenland S. Tests for interaction in epidemiologic studies: a review and a study of power. *Stat Med*. 1983;2:243–251.
20. Breslow NE, Day NE. *Statistical Methods in Cancer Research II: The Design and Analysis of Cohort Studies*. Lyon: International Agency for Research on Cancer; 1987.
21. Greenland S. Basic problems in interaction assessment. *Enivon Health Perspect*. 1993;101(Suppl 4):59–66.
22. Greenland S. Additive risk versus additive relative-risk models. *Epidemiology*. 1993;4:32–36.
23. Greenland S. Multivariate estimation of exposure-specific incidence from case-control studies. *J Chronic Dis*. 1981;34:445–453.
24. Greenland S. Introduction to regression modeling. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008:418–455.
25. Greenland S. Summarization, smoothing, and inference. *Scand J Soc Med*. 1993;21:227–232.
26. Greenland S. Smoothing observational data: a philosophy and implementation for the health sciences. *Int Stat Rev*. 2006;74:31–46.
27. Greenland S. Variable selection and shrinkage in the control of multiple confounders. *Am J Epidemiol*. 2008;167:523–529.
28. Friedman JH, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting. *Ann Stat*. 2000;28:337–374.
29. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer; 2001.
30. Ridgeway G, McCaffrey DF. Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data. *Stat Sci*. 2007;22:540–543.

31. Davison AC, Hinkley DV. Bootstrap methods and their application. New York: Cambridge; 1997.
32. Greenland S. The effect of misclassification in the presence of covariates. *Am J Epidemiol*. 1980;112:564–569.
33. Greenland S. When should epidemiologic regressions use random coefficients? *Biometrics*. 2000;56:915–921.
34. Greenland S, Lash TL. Bias analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008:345–380.
35. Greenland S. Multiple-bias modeling for analysis of observational data (with discussion). *J R Stat Soc Ser A*. 2005;168:267–308.
36. Greenland S. Introduction to regression models. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008:381–417.