

Outcome-wide Epidemiology

Tyler J. VanderWeele

Abstract: The author proposes that epidemiologic studies should more often assess the associations of a single exposure with multiple outcomes simultaneously. Such “outcome-wide epidemiology” will be especially important for exposures that may be beneficial for some outcomes but harmful for others. Outcome-wide epidemiology may also be helpful in prioritizing public health recommendations. Methodologically, the conduct of outcome-wide epidemiology will generally be more straightforward than recent proposals for exposure-wide epidemiologic studies, in which the associations between a single outcome and many exposures are assessed simultaneously. Such exposure-wide studies are likely to be subject to numerous biases because of the inability to make simultaneous confounding control and because exposures are likely to affect, and mediate the effects of, other exposures. These problems simplify considerably in an outcome-wide approach when a single exposure is being considered. Moreover, outcome-wide approaches will generally be more useful than exposure-wide approaches in shaping public health recommendations.

(*Epidemiology* 2017;28: 399–402)

Due in part to the success of genome-wide association studies, there has been a recent suggestion that the research community begin to move toward an “exposure-wide epidemiology,”¹ in which associations between an outcome and many exposures—possibly very many exposures—are assessed simultaneously.¹ I would like to argue here that, due to the nature of confounding, attempts at “exposure-wide epidemiologic” studies are likely to be plagued by biases, but that instead an “outcome-wide epidemiology” may in fact be more feasible, and that, moreover, such outcome-wide epidemiology will likely also be more relevant for public health recommendations.

LIMITS OF EXPOSURE-WIDE EPIDEMIOLOGY

The notion of an exposure-wide epidemiologic study is that a researcher could select a specific outcome, regress

it upon a wide range of different exposures, assess which relationships are most substantial, and for which there is the strongest statistical evidence of an association, and, provided appropriate control is made for multiple testing, thereby potentially gain insight into the underlying causes of the disease or outcome under study. This approach has effectively been what has been used in genome-wide association studies,^{1,2} and these have now yielded thousands of replicated associations between genetic variants and various diseases.^{3,4} Could we not then pursue something similar within epidemiology more broadly using environmental, social, and behavioral exposures as well?

The difference between genetic exposures and almost all others, and the difference that arguably creates problems for an exposure-wide epidemiology, lies in the nature of confounding. In a genome-wide association study, although hundreds of thousands of variants are examined, it is often thought to be the case that, subject to control for population stratification (often done say by principal components analysis adjustment strategies), the association between the variant and the outcome is roughly unconfounded.² While a particular variant may serve as a proxy for the true effect of another, it is the case that once the genome is fixed, each variant is acting on the outcome, possibly in conjunction with, but not by altering the value of, any other variant. This is manifestly not the case with environmental, behavioral, and social exposures, wherein one exposure is likely to affect many others downstream. Each exposure will thus likely require a distinct set of other variables to control for confounding, with the confounding variables for a particular exposure consisting only of other exposures that are temporally before it. If we include all of our exposures in the model and some of these are downstream from others, then the downstream exposures will likely mediate, and potentially block, the effects of prior exposure.

This creates two potential problems. First, for each exposure, the association estimate will, at best, represent the direct effect of the exposure not through any of the other exposures in the model downstream of it. If there are numerous subsequent exposures that mediate the effect of the prior exposure then the importance of the prior exposure (in terms of its overall influence on the outcome) might be severely misrepresented.^{5,6} Second, it is now well documented in the methodology literature that if control is made for mediating variables on pathways from exposure to outcome, then any unmeasured common cause of the mediating variable and the

Submitted 1 July 2016; accepted 31 January 2017.

From the Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA.

Supported by NIH Grant R01 ES017876.

The author reports no conflicts of interest.

Correspondence: Tyler J. VanderWeele, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115. E-mail: tvanderw@hsph.harvard.edu.

Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 1044-3983/17/2803-0399

DOI: 10.1097/EDE.0000000000000641

outcome can induce bias; spurious associations between exposure and outcome can be generated even if the exposure has no effect on the outcome whatsoever. This problem is sometimes referred to in the literature as one of “collider stratification bias.”^{7,8} When considering multiple exposures simultaneously, the likelihood of such biases is substantial.

EXAMPLES AND CHALLENGES IN CONFOUNDING CONTROL

Let us illustrate with some examples. Consider the first problem concerning mediating variables: Suppose we had an exposure-wide epidemiologic study examining self-rated health as the outcome. We might consider race, age, educational attainment, various health behaviors, and so on, as possible exposures. How would the exposure “race” fare in such an exposure-wide regression model? Analyses using the National Longitudinal Survey of Youth data indicate that, after control for a measure of educational test scores, the gap between black and white men in self-reported health effectively disappears entirely.^{9,10} Are we thus to conclude that race is irrelevant here as an exposure? This would be misleading as race affects educational opportunities that are themselves arguably on the pathway from race to health. The exposure-wide approach, by controlling for downstream factors, does not capture the overall importance of the variable.

One might then take the position that such exposure-wide approaches can be useful, but need to be interpreted carefully as only the direct effect not through other exposures. However, even this is potentially problematic because of the potential for collider stratification bias discussed above. Consider an exposure-wide epidemiologic study examining infant mortality as an outcome with maternal smoking, age, and birthweight among the exposures. Numerous studies have documented a seemingly protective association between maternal smoking and infant mortality for low-birthweight infants.^{11,12} One explanation for these paradoxical associations is again collider stratification bias, an unmeasured common cause of low birthweight and infant mortality (e.g., malnutrition or a birth defect).^{13,14} Such an unmeasured common cause essentially sets up an unfair comparison between smoking and nonsmoking mothers with low-birthweight infants: for the nonsmoking mothers who have low-birthweight infants, smoking cannot be the cause of low birthweight so the cause might well instead have been something like malnutrition or a birth defect, the consequences of which for infant mortality are worse than smoking.

An exposure-wide epidemiologic study which ignored such potential biases might come away with the conclusion that among low-birthweight infants, maternal smoking is protective for infant mortality. While someone might respond that it is simply the case that such biases must be thought through in exposure-wide studies, just as in regular epidemiologic investigation, the number of potential instances of such biases that must be considered when dozens, or hundreds, of exposures

are considered simultaneously is mind-boggling. The field of epidemiology currently struggles with these issues in studies of a single exposure. It is arguably not reasonable then to think that we could do this adequately when numerous exposures are considered at once. Moreover, even if we could, we would still only be obtaining direct effects as above.

The typical epidemiologic approach toward confounding control when examining the effect on an outcome of a single exposure at a specific point in time is to consider all possible temporally prior variables that might affect the exposure or the outcome, including possibly prior values of the exposure and the outcome.^{15,16} While precise confounding control strategies do vary,^{17,18} there is consensus that, if the total effect of the exposure on the outcome is desired, then adjustment should not be made for variables that might be affected by the exposure. The implication of this, as indicated above, is that for each individual exposure, we will likely need a distinct set of confounding variables. We cannot make the decision about confounding for all variables at once when we are supposedly examining the effects of multiple exposures. A single regression model will not suffice; nor will simply looking at each bivariate association one at a time, as in genome-wide studies. This arguably creates difficulties for a simple approach to exposure-wide epidemiology.

OUTCOME-WIDE EPIDEMIOLOGY

Consider instead now what one might call an “outcome-wide epidemiologic study” in which only a single exposure is under study but its effects on multiple outcomes are being considered simultaneously. In such a design, we could potentially attempt to control for confounding for the effect of the exposure on all outcomes simultaneously. This could be done by attempting to control for all variables temporally prior to the exposure that might affect the exposure (and possibly any one or more of the outcomes). If we simply control for all variables temporally prior to the exposure and include this same set of covariates as control variables for each outcome, then the confounding control decision effectively only has to be made once. While some variables may confound the relationship between the exposure and one outcome, but not another, we do still have the option, unlike in the exposure-wide epidemiologic setting, of simply controlling for all variables prior to the exposure. With the exposure fixed, the set of all variables temporally before the exposure stays the same even when we change the outcome. With the outcome fixed, the set of variables temporally before the exposure changes as we change the exposure.

To illustrate the difference further, consider the causal diagram^{19–21} in the Figure with five exposures, E_1 – E_5 , some baseline covariates C , an outcome Y , and a single unmeasured common cause U of E_4 and Y . In an exposure-wide study which regressed Y on all five exposures along with the covariates C , we would obtain biased estimates of the direct effect of E_4 on Y because of the unmeasured confounder U , but the regression would also give biased estimates of the direct effects of E_1 , E_2 ,

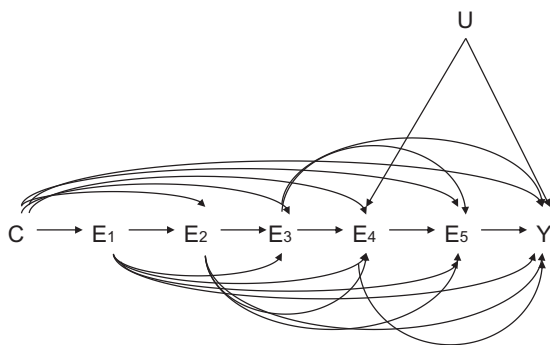


FIGURE. Causal diagram illustrating the confounding challenges that can arise in trying to assess the effects of multiple exposures (E_1 – E_5) on a single outcome (Y) in the presence of measured (C) and unmeasured (U) confounding variables.

and E_3 on Y because of collider stratification^{7,8} as discussed above: E_4 is common effect of U and each of E_1 , E_2 , and E_3 so controlling for it in the regression will introduce spurious associations between each of E_1 , E_2 , and E_3 and Y because of U . Thus, in our exposure-wide epidemiologic study, the causal effect estimates (even acknowledging these are at best direct effects) of four out of our five exposures will be biased by a single unmeasured confounder. In contrast, an outcome-wide epidemiologic study that simply focused on E_3 would obtain unbiased estimates of the effect of E_3 on the outcomes E_4 , E_5 , and Y . Likewise, an outcome-wide epidemiologic study of the effects of E_1 or of E_2 on everything subsequent would be unbiased. We would still have bias in an outcome-wide analysis of the effects of E_4 , because of confounding, but as seen above, for many exposures the outcome-wide approach would be free of bias. A single unmeasured confounder of a particular exposure–outcome relationship introduces bias in an outcome-wide analysis for that particular exposure; but a single unmeasured confounder in an exposure-wide analysis generates bias for that exposure and also all prior exposures that affect that exposure. The number of potential confounding relationships and biases are much broader in an exposure-wide analysis than in an outcome-wide analysis.

PUBLIC HEALTH IMPORTANCE OF OUTCOME-WIDE EPIDEMIOLOGY

However, perhaps an even more powerful argument for outcome-wide epidemiology than the relative simplicity of confounder control concerns its potential use for public health recommendations. Epidemiology has traditionally examined exposure–outcome relationships one at a time. This is arguably still the soundest approach for assessing evidence for causality. However, many exposures are likely to affect a variety of outcomes. When effects are nearly universally beneficial, as with say low-impact exercise on health, public health recommendations are straightforward. But other exposures may affect different health or life outcomes in divergent directions: some beneficial, some harmful. Moderate drinking may have

some benefit for all-cause mortality, but also increases the likelihood of accidents and liver cirrhosis.^{22–24} For postmenopausal women, studies indicate that estrogen plus progestin increases the likelihood of coronary heart disease, breast cancer, and stroke but decreases the risk of colorectal cancer and hip fractures.^{25,26} In making public health recommendations concerning a particular exposure that may be favorable for some outcomes and unfavorable for others, it would arguably be helpful to examine associations between that exposure and many different outcomes simultaneously, rather than just one at a time across studies. An outcome-wide epidemiologic study looking at a broad range of health and life outcomes could provide considerable insight into what the trade-offs might be. Such analyses could provide a more nuanced set of recommendations and potentially be helpful for individual decision-making as well. The analyses could also potentially be stratified by gender and age, or by race, or, when sufficient data are available, perhaps also by other variables as well. Even when exposures are beneficial across a wide range of outcomes, such outcome-wide epidemiologic studies might be helpful in prioritizing public health recommendations by identifying those exposures with the largest relative effect sizes across numerous outcomes. Large cohort studies with extensive covariate information, long-term follow-up, and data on a wide range of outcomes would of course be extremely useful both for the purposes of confounding control and for examining numerous outcomes, including rarer outcomes, simultaneously.

DISCUSSION

I do not want to dismiss the potential utility of exposure-wide epidemiologic studies entirely. A single regression model for a given outcome using many exposures may be valuable for predictive purposes. With cohort data for which repeated measures of exposures are available, another approach might be to examine a single outcome at the end of follow-up and fit a series of regressions, each of which controls for all exposures simultaneously in one wave but then also includes a single subsequent exposure—one per regression—from the next wave.^{15,16} One might refer to this as a “confounder-lagged exposure-wide epidemiologic design.” See Betancourt et al.²⁷ for a recent example in which both this and an outcome-wide approach were simultaneously used in the study of mental health outcome for war-affected youth. Other variations that attempt to straightforwardly automate the confounder selection process for exposure-wide epidemiologic studies containing numerous exposures might also be considered.

I also do not want to minimize the potential biases that may still be present in an outcome-wide epidemiologic study. Confounding control decisions, even for a single exposure, can be challenging with numerous subtleties^{16–21,28–31} and temporal ordering may not always be clear; issues of measurement error and selection bias must likewise be evaluated carefully²¹; and careful modeling including potential interactions between the exposure and covariates may be necessary.

Focusing on a single exposure certainly does not sweep all problems away. However, I would still maintain that, from the perspective of public health recommendations, and individual decision-making, the outcome-wide approach will be of more direct and immediate benefit. For exposures with both beneficial and detrimental effects, epidemiologists should perhaps be employing such approaches routinely.

REFERENCES

- Ioannidis JP. Exposure-wide epidemiology: revisiting Bradford Hill. *Stat Med*. 2016;35:1749–1762.
- Hunter DJ. Lessons from genome-wide association studies for epidemiology. *Epidemiology*. 2012;23:363–367.
- Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014;42(Database issue):D1001–D1006.
- GWAS Catalog. Available at: <https://www.ebi.ac.uk/gwas/> Accessed 1 July 2016.
- MacKinnon D. *Statistical Mediation Analysis*. New York, NY: Routledge; 2008.
- VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York, NY: Oxford University Press; 2015.
- Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
- Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39:417–420.
- Fryer R. Racial inequality in the 21st century: the declining significance of discrimination. In: Orley A, David C, eds. *Handbook of Labor Economics*, vol. 4, part B. Amsterdam, Netherlands: Elsevier; 2011:855–971.
- Neal DA, Johnson WR. The role of premarket factors in black-white wage differences. *J Polit Econ*. 104: 869–895.
- Yerushalmy J. The relationship of parents' cigarette smoking to outcome of pregnancy—implications as to the problem of inferring causation from observed associations. *Am J Epidemiol*. 1971;93:443–456.
- Wilcox AJ. Birth weight and perinatal mortality: the effect of maternal smoking. *Am J Epidemiol*. 1993;137:1098–1104.
- Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight “paradox” uncovered? *Am J Epidemiol*. 2006;164:1115–1120.
- VanderWeele TJ, Mumford SL, Schisterman EF. Conditioning on intermediates in perinatal epidemiology. *Epidemiology*. 2012;23:1–9.
- Hernán MA. Counterpoint: epidemiology to guide decision-making: moving away from practice-free research. *Am J Epidemiol*. 2015;182:834–839.
- VanderWeele TJ, Jackson JW, Li S. Causal inference and longitudinal data: a case study of religion and mental health. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51:1457–1466.
- Sauer BC, Brookhart MA, Roy J, VanderWeele TJ. *Covariate selection. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide*, Chapter 7. Agency for Healthcare Research and Quality: Rockville, MD; 2013:93–108.
- VanderWeele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics*. 2011;67:1406–1413.
- Pearl J. *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press; 2009.
- Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8:70.
- Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2008.
- Di Castelnuovo A, Costanzo S, Bagnardi V, Donati MB, Iacoviello L, de Gaetano G. Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies. *Arch Intern Med*. 2006;166:2437–2445.
- Rehm J, Taylor B, Mohapatra S, et al. Alcohol as a risk factor for liver cirrhosis: a systematic review and meta-analysis. *Drug Alcohol Rev*. 2010;29:437–445.
- Cherpitel CJ. Alcohol and injuries: a review of international emergency room studies since 1995. *Drug Alcohol Rev*. 2007;26:201–214.
- The Women's Health Initiative Study Group. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative. *JAMA*. 2002;288:321–333.
- Manson JE, Hsia J, Johnson KC, et al.; Women's Health Initiative Investigators. Estrogen plus progestin and the risk of coronary heart disease. *N Engl J Med*. 2003;349:523–534.
- Betancourt TS, Gilman SE, Brennan RT, Zahn I, VanderWeele TJ. Identifying priorities for mental health interventions in war-affected youth: a longitudinal study. *Pediatrics*. 2015;136:e344–e350.
- Pearl J. *On a Class of Bias-Amplifying Variables That Endanger Effect Estimates*. Corvallis, OR: Association for Uncertainty in Artificial Intelligence; 2010.
- Myers JA, Rassen JA, Gagne JJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. *Am J Epidemiol*. 2011;174:1213–1222.
- Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*. 2003;14:300–306.
- Ding P, Miratrix LW. To adjust or not to adjust? Sensitivity analysis of M-Bias and Butterfly-Bias (with comments). *J Causal Inference*. 2015;3:41–57.