Title
Theory and methodology: essential tools that can become dangerous belief systems

Journal
European Journal of Epidemiology, 33(5)

ISSN
0393-2990

Authors
Greenland, Sander
Jewell, Nicholas Patrick
Mansournia, Mohammad Ali

Publication Date
2018-05-01

DOI
10.1007/s10654-018-0395-7

Peer reviewed

HHS Public Access

Author manuscript

Eur J Epidemiol. Author manuscript; available in PMC 2020 September 04.

Published in final edited form as:

Eur J Epidemiol. 2018 May ; 33(5): 503–506. doi:10.1007/s10654-018-0395-7.

Theory and methodology: essential tools that can become dangerous belief systems

Sander Greenland[1,2], **Nicholas Patrick Jewell**[3,4], **Mohammad Ali Mansournia**[5]

[1]Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles, CA, USA

[2]Department of Statistics, College of Letters and Science, University of California, Los Angeles, CA, USA

[3]Division of Biostatistics, School of Public Health, University of California, Berkeley, CA, USA

[4]Department of Statistics, University of California, Berkeley, CA, USA

[5]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, PO Box 14155-6446, Tehran, Iran

Abstract

We thank Dr. Karp for his interest [1] in our paper [2]. We agree on some points, but our theoretical description differs from his in ways leading to important divergences for teaching and practice. We also see a danger of overextending abstract theory (with its inevitable and extensive simplifications) into practice [3], especially when the practical questions are causal but the theory applied lacks an explicit, sound longitudinal causal model to address these questions. As we will explain, a defect in the "study base" theory Dr. Karp adopts as a foundational belief system is that it takes as a foundation a parameter affected by baseline risk factors—including exposure when that has effects on follow-up or disease. It consequently leads to biases and misconceptions of the sort documented elsewhere [4, 5] and below, which require a coherent theory of longitudinal causality to address. Our divergence from Dr. Karp thus raises the issue of the role of theory and methods in research, although matching serves to illustrate our points in a familiar epidemiologic context.

Keywords

Case–control studies; Causal inference; Confounding; Epidemiological research

Moving on from old theories to eradicate misconceptions

Dr. Karp discusses the "trohoc fallacy," which involves thinking that case–control studies are based on comparing cases to noncases, rather than recognizing that as with cohort studies they compare exposure groups. We are not sure why he raised that issue; we didn't discuss it since that fallacy, while common in the 1970s, was thoroughly discredited by the 1980s.

mansournia_ma@yahoo.com.

True, the fallacy still appears occasionally in methods papers [e.g., 6] and is embedded in the archaic view of case–control studies as inherently retrospective (which confuses the order of etiologic and selection events with order of data recording). But credible methodologic writings have long taken pains to emphasize that the causally relevant target comparisons are across exposures, not outcome groups (regardless of whether there is oversampling of cases relative to noncases) [7, 8]. In fact, the idea that the case–control design is best viewed as an efficient sampling strategy for study of uncommon outcomes has been around for generations and remains part of standard theory [9, Ch. 8]. This conceptualization includes dynamic sampling of both cases and time at risk (risk-set or density control sampling) to permit direct estimation of hazard (incidence-rate) ratios under a proportional-hazards model, an idea that dates back at least to the early 1960s [10]. So we see no disagreement here between Dr. Karp's position and ours.

Unlike Dr. Karp, however, we long ago discarded the idea that person-time and hazard rates ("incidence densities") are fundamental or provide a sound foundation for design or analysis of studies of causality [11]. Rather than being modern in the sense of informed by recent developments, we find the "study base"/"person moments" view (which was in vogue in the 1980s) to be dated and potentially misleading for both teaching and practice. The underlying issue has long been recognized in warnings to not adjust for post-randomization covariates [12] (now modified to: adjust using only proper longitudinal causal models [13, 14]). The problem is that person-time denominators are in fact po0st-treatment covariates affected by variables that affect follow-up or risk (i.e., disease, censoring, and competing-risk processes). Those variables will usually include the very exposures under study, even if those have no effect on the study disease (e.g., if they have an effect on follow-up or on general mortality).

As a consequence, contrasts of person-time rates or 'incidence densities' are subject to artefacts such as non-collapsibility and time-varying confounding by fixed covariates (such as biological sex)—even under randomization [14–17]; those artefacts can be severe enough to produce associations in the opposite direction of effects ("crossing hazards"). Because of these and related defects in person-time "study-base" theory, modern longitudinal causality theory is instead founded on concepts of risk and survival times in cohorts, with rates and open (dynamic) populations serving as important statistical concepts to be used with due caution [9, Ch. 3; 14, 18].

Matching has varying rationales, forms, costs, and benefits

On the practical side, matching of cohorts at baseline is an effective tool in the control of confounding by variables fully determined at baseline; and, with the spread of propensity-score matching, it has become more common than ever for the formation of study cohorts. Variables whose values or trajectories known at baseline include several of high concern in typical studies, notably genetic factors (which are fixed) and age (time-varying but exogenous, being fully determined by date of birth). Dr. Karp seems to overlook that, even if a variable is time varying and thus a trajectory, matching on its baseline value can help reduce trajectory imbalances among the matched subcohorts, and thus improve efficiency of subsequent longitudinal analyses (although, as with simple matching [19], improvement is

not guaranteed). For example, baseline matching on the category “vegetarian yes/no” will not control much of the complex and time-varying confounding by diet, but may well reduce bias and aid analytic control to the extent such partial matching brings the compared cohorts closer in dietary distributions.

The framework of partial matching [9, p. 182; 20] is more nuanced than that of all-or-none matching (which most matching discussions presume is the only choice). No practical account of matching is complete without this framework—especially since partial matching is often the most efficient strategy in both practical and statistical terms [21]. It is also ubiquitous in practice: Most matching on continuous variables uses categories or calipers and is thus partial, leaving some imbalance (and thus potential confounding) within the categories. This is a common problem for age matching in broad categories, and requires special analytic adjustments to address [9, pp. 434–435]; e.g., simply tossing age in a model will not remove the residual confounding [20, 22].

Matching and research reality

Most discussions of matching including ours [2], and indeed most mathematical analyses of epidemiologic methods, can be criticized as oversimplified relative to the actual options and realities of epidemiologic research. It is thus important to understand that, like all theoretical analyses, they serve as warnings about what can go wrong even in simple cases, but going beyond such warnings requires caution in view of the hazards of extrapolation. Also essential is some attention to likely magnitudes of problems in light of real-data experience [23]. We thus regard as potentially misleading Dr. Karp’s comment that

> regardless of which type of ‘controls’ is used, and regardless of whether case–control matching is used, adjustment for “confounding by variables that are difficult to measure” can only be done in the statistical ‘analysis’—by conditioning on such ‘variables,’ possibly operationally represented by their nominal-scale proxies (such as sibship status or neighborhood).

This comment (especially the “only”) ignores the multiplicity of bias sources, their interactions, and their net consequences. For example, when the association of exposure with the matching factors is modest, we have noticed that the unmatched (unadjusted) estimator from the matched data usually suffers only a minor amount of bias toward the null from breaking the matches. Then too, if the matching factors are strong confounders by virtue of a strong effect on risk (as is age, typically) there can be much less bias in that estimator than would be found in the unadjusted estimator from an unmatched version of the study. In that setting we again have noticed that case–control matching results in much smaller net bias than the original confounding by the matching factors, and so accomplishes a partial adjustment without an analytic step. Thus, while adjustment for matching factors is completed by their analytic control, and we have always recommended that be done [2, 9, p. 176], in our experience the net impact of skipping that adjustment has usually been small unless the age categories are too broad.

We also think important qualifications are needed for Dr. Karp’s statement that “In the context of the need for such an adjustment for confounding control, matching may merely

help enhance the efficiency of the analysis (and possibly reduce the risk/extent of the 'sparse-data bias.')." For example, common adjustments (e.g., conditioning on the matched sets) may generate or aggravate sparse-data bias [24–26], a bias that is away from the null that may exceed the bias in the unmatched estimator from either a matched or unmatched design. And to the extent matching forces us to control factors that could have been ignored without it, is can increase sparse-data bias. In either case, simple adjustments for sparse-data bias are available in standard software [26–29] and should be deployed more often.

Regarding general cost efficiency, neighborhood matching may drastically reduce the time and cost of locating and visiting population controls; in contrast, demanding close matches on personal factors may drastically increase time and cost. And if efficiency is the only concern, modified matching (e.g., partial matching, marginal matching, countermatching) may be superior to traditional all-or-none matching for both statistical and practical reasons [20, 30, 31], although we again caution that propensity-score matching may introduce artefacts [32]. Finally, there is a continuum of partial-adjustment options between full and no adjustment, with full adjustment not always the best option for estimation accuracy [33].

Conclusion

Health and social sciences are intractable in full complexity; as a consequence, theory and methodology provide only guidelines and heuristics [3, 23, 34]. Sometimes these guidelines may always benefit practice, and so no harm is done by taking them as natural laws; but more often, transformation of methodologic theory from a toolkit into a metaphysical belief system can blind us to its important exceptions and failings (as will inevitably arise as research expands). The only safe general advice would then seem to be: For every guideline offered or followed, understand the assumptions and simplifications in the theory used to derive that guideline, so that its limits and breakdowns can be recognized. And be prepared to revise or replace established methodologic theories when they are superseded by theories that handle old problems just as well and also solve problems with the old theories. This is precisely what has happened as causal (structural) models and causal diagrams have replaced traditional population models as core concepts for causal inference [14, 35–37], in the process demoting rates and population-time to an intermediate (although still vital) statistical role in linking observed associations to causal effects [9].

References

1. Karp I Toward eradicating misconceptions on matching in etiological studies. Eur J Epidemiol. 2018 10.1007/s10654-018-0376-x.
2. Mansournia MA, Jewell NP, Greenland S. Case–control matching: effects, misconceptions, and recommendations. Eur J Epidemiol. 2018;33:5–14. [PubMed: 29101596]
3. Greenland SFor and against methodology: some perspectives on recent causal and statistical inference debates. Eur J Epidemiol. 2017;32(1):3–20. [PubMed: 28220361]
4. Greenland S Confounding of incidence density ratio in case–control studies. Epidemiology. 2013;24:624–5. [PubMed: 23732742]
5. Pang M, Schuster T. Confounding of incidence density ratio in case–control studies. Epidemiology. 2013;24:625–7. [PubMed: 23732743]

6. Allen AS, Glen A, Satten GA. Control for confounding in case–control studies using the stratification score, a retrospective balancing score. Am J Epidemiol. 2011;173:752–60. [PubMed: 21402731]
7. Breslow NE, Day NE. Statistical methods in cancer research. Vol I: the analysis of case–control data. Lyon: IARC; 1980.
8. Rothman KJ. Modern epidemiology. Boston: Little, Brown; 1986.
9. Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2008.
10. Sheehe PR. Dynamic risk analysis in retrospective matched-pair-studies of disease. Biometrics. 1962;18:323–41.
11. Greenland S Cohorts versus dynamic populations: a dissenting view. J Chronic Dis. 1986;39:565–6. [PubMed: 3722320]
12. Cox DR. The planning of experiments. New York: Wiley; 1958.
13. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. BMJ. 2017;359:j4587. [PubMed: 29038130]
14. Hernán MA, Robins JM. Causal inference. New York: Chapman and Hall; 2018.
15. Greenland S Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. Epidemiology. 1996;7:498–501. [PubMed: 8862980]
16. Hernán MA. The hazards of hazard ratios. Epidemiology. 2009;20:13–5.
17. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. Int J Epidemiol. 2013;42:860–9. [PubMed: 23918854]
18. Cole SR, Hudgens MG, Brookhart MA, Westreich D. Risk. Am J Epidemiol. 2015;181:246–50. [PubMed: 25660080]
19. Greenland S, Morgenstern H. Matching and efficiency in cohort studies. Am J Epidemiol. 1990;131:151–9. [PubMed: 2293747]
20. Partial Greenland S. and marginal matching in case–control studies In: Moolgavkar SH, Prentice RL, editors. Modern statistical methods in chronic disease epidemiology. New York: Wiley; 1986 p. 35–49.
21. Stürmer T, Brenner H. Degree of matching and gain in power and efficiency in case–control studies. Epidemiology. 2001;12:101–8. [PubMed: 11138803]
22. Greenland S Re: “Estimating relative risk functions in case–control studies using a nonparametric logistic regression”. Am J Epidemiol. 1997;146:883–4. [PubMed: 9384210]
23. Intuitions Greenland S., simulations, theorems: the role and limits of methodology (invited commentary). Epidemiology. 2012;23:440–2. [PubMed: 22475828]
24. Greenland S Small-sample bias and corrections for conditional maximum-likelihood odds-ratio estimators. Biostatistics. 2000;1:113–22. [PubMed: 12933529]
25. Greenland S, Mansournia MA. Penalization, bias reduction, and default priors in logistic and related categorical and survival regressions. Stat Med. 2015;34:3133–43. [PubMed: 26011599]
26. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. BMJ. 2016;352:i1981. [PubMed: 27121591]
27. Sullivan S, Greenland S. Bayesian regression in SAS software. Int J Epidemiol. 2013;42:308–17. [PubMed: 23230299]
28. Discacciati A, Orsini N, Greenland S. Approximate Bayesian logistic regression via penalized likelihood by data augmentation. Stata J. 2015;15(3):712–36.
29. Mansournia MA, Geroldinger A, Greenland S, Heinze G. Separation in logistic regression—causes, consequences, and control. Am J Epidemiol. 2018;187:864–70. 10.1093/aje/kwx299. [PubMed: 29020135]
30. Stürmer T, Brenner H. Flexible matching strategies to increase power and efficiency to detect and estimate gene-environment interactions in case–control studies. Am J Epidemiol. 2002;155:593–602. [PubMed: 11914186]
31. Langholz B, Clayton D. Sampling strategies in nested case–control studies. Environ Health Perspect. 1994;102(Suppl 8):47–51.

32. Mansson R, Joffe MM, Sun W, Hennessy S. On the estimation and use of propensity scores in case–control and case–cohort studies. Am J Epidemiol. 2007;166:332–9. [PubMed: 17504780]

33. Kalish LA. Reducing mean squared error in the analysis of pair-matched case–control studies. Biometrics. 1990;46:493–9. [PubMed: 2364134]

34. Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. Int J Epidemiol. 2016;45:1776–86. [PubMed: 26800751]

35. Pearl J Causality: models, reasoning and inference. 2nd ed. Cambridge: Cambridge University Press; 2009.

36. VanderWeele TJ. Explanation in causal inference: methods for mediation and interaction. New York: Oxford University Press; 2015.

37. Pearl J, Glymour M, Jewell NP. Causal inference in statistics: a primer. New York: Wiley; 2017.