Social Science & Medicine xxx (2016) 1–8

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Causal inference challenges in social epidemiology: Bias, specificity, and imagination

M. Maria Glymour [a, *], Kara E. Rudolph [b, c]

[a] Department of Epidemiology and Biostatistics, University of California, San Francisco, USA
[b] Center for Health and Community, University of California, San Francisco, USA
[c] School of Public Health, University of California, Berkeley, USA

ARTICLE INFO

Article history:
Received 9 October 2015
Received in revised form
26 July 2016
Accepted 31 July 2016
Available online xxx

Keywords:
Population health
Causal inference
Research methods
Propensity scores
Fixed effects models
Instrumental variables
Social epidemiology

> “Efficient scientific iteration evidently requires unhampered feedback. …. In any feedback loop it is, of course, the error signal - for example, the discrepancy between what tentative theory suggests should be so and what practice says is so - that can produce learning. The good scientist must have the flexibility and courage to seek out, recognize, and exploit such errors - especially his own …. he must not be like Pygmalion and fall in love with his model” (Box, 1976)

Social epidemiology is closely linked with an imperative for public health action to eliminate health inequalities. Interest in causal inference has therefore long motivated the field, even when the language of causality is not explicitly invoked. In recent years, social epidemiologists overtly adopted a language and tool set focused on supporting causal inferences from observational evidence. A handful of disappointing trials showing no effects of social interventions on primary health outcomes (Glass et al., 2004; Writing Committee for the Enrichd Investigators, 2003) intensified the field's enthusiasm for research methods that promised to strengthen causal inference (Berkman, 2009). The null results of the small number of trials in social epidemiology coincided roughly with a series of more publicly discussed setbacks for other domains of observational epidemiology, for example in the unexpectedly adverse effects observed in the beta-carotene and hormone replacement therapy trials (Smith et al., 2001; Hulley et al., 1998; Manson et al., 2003; Omenn et al., 1996). The combination fostered a reappraisal of the methodological toolkit of observational epidemiology, with many researchers, including social epidemiologists, welcoming new methods to improve the quality of evidence for causal inference (Hernán et al., 2008; Oakes, 2004; Oakes et al., 2006).

Given that social epidemiologists have historically asked causal questions (even when they cloaked their conclusions in associational language), how does prior work differ from the recent outpouring of methods and writing that explicitly invokes causal inference goals? The most marked divide between modern causal inference approaches and traditional approaches is not the statistical machinery, but rather that a causal inference approach begins by articulating the assumptions necessary to support causal inferences (e.g., that treatment assignment is independent of the potential outcomes, conditional on included covariates). Once these assumptions are stated clearly, the magnitude of the challenge to design and analyze a study that satisfies or relaxes these assumptions becomes apparent. Alongside greater transparency of assumptions and goals, the field is also taking advantages of recent statistical developments and borrowing methods from other areas of research. In this commentary, we discuss our views on the priorities for strengthening and enriching causal inferences in social epidemiology. We organize our conceptualization into three related challenges: bias, specificity, and imagination.

1. Bias

The challenge of bias – systematically incorrect estimates of causal effects – has been the predominant focus of much causal

* Corresponding author. Department of Epidemiology and Biostatistics, University of California, San Francisco 550 16th Street, San Francisco, CA 94158, USA.
E-mail address: mglymour@epi.ucsf.edu (M.M. Glymour).

http://dx.doi.org/10.1016/j.socscimed.2016.07.045
0277-9536/

Please cite this article in press as: Glymour, M.M., Rudolph, K.E., Causal inference challenges in social epidemiology: Bias, specificity, and imagination, Social Science & Medicine (2016), http://dx.doi.org/10.1016/j.socscimed.2016.07.045

inference research. This concern has largely focused on internal validity (drawing correct causal inferences about the sample participating in the study), rather than on external validity (drawing inferences that describe causal relations in a target population other than those originally studied). We return to external validity as it relates to providing evidence with sufficient precision to guide trial design in the Specificity section.

There is certainly good reason for epidemiologists to worry about internal validity. RCTs are expected to have high internal validity because the randomization process balances potential confounders across treatment groups. In nonrandomized observational studies, however, such balance is not inherent in the study design. The disappointing RCTs in social and other domains of epidemiology led to impassioned debate about the usefulness of observational research tools to anticipate trial results (Smith et al., 2001; Taubes et al., 1995). This debate holds special weight in social epidemiology because many social factors will probably never be subject to an ideal RCT. Thus, if the field concludes that observational evidence cannot be trusted, vast swathes of social epidemiology must be abandoned.

However, several studies have directly compared findings from rigorously designed observational studies with results of RCTs addressing comparable research questions (Anglemyer et al., 2014; Cook et al., 2008; Dehejia et al., 1999; Hernán et al., 2008; Ioannidis et al., 2001). Findings from these studies suggest that with appropriate design and analysis methods, observational research can deliver findings in line with RCT results. However, these methods are still rare in social epidemiology. We remark on a handful methodological approaches for strengthening internal validity that have received increasing emphasis in the last decade in social epidemiology (including but not limited to those commonly encountered in the field of causal inference): (1) propensity score approaches, including inverse probability weighting; (2) approaches based on within-person changes in exposure, including difference-in-difference methods; (3) pseudo-randomization approaches leveraging quasi- or natural experiments; and (4) sensitivity analyses for unobserved confounding and measurement error. A table summarizing some strengths and weaknesses, providing key citations, and examples is available at http://bit.ly/2a9hjNR. Although we do not focus here on aspects of study design and implementation, these are clearly essential for achieving internal validity. Often decisions made at the design stage – e.g., regarding sampling, measurement, and follow-up protocols – limit the types of analyses that will be possible.

1.1. Modeling exposure: propensity score methods

Confounding arises from common prior causes of the exposure and outcome, and so can be addressed by modeling either the outcome (as in traditional regression adjustment), modeling the exposure, or modeling both the exposure and outcome. Propensity score methods include a set of approaches based on modeling the exposure. A propensity score is defined as the probability of having the exposure of interest compared to a reference exposure (Rosenbaum et al., 1983). For example, in the simple case of a binary exposure, the propensity score might be estimated using a logistic regression model, in which the binary exposure variable is a function of potential confounding variables.

Propensity score methods are designed to balance the distribution of measured covariates (potential confounders) between the exposed and unexposed group. If there are no unmeasured confounders, this balancing of covariates mimics an RCT. The most familiar propensity score method may be 1:1 nearest neighbor propensity score matching, in which each exposed individual is compared to an unexposed individual with a similar propensity score. Numerous other propensity score approaches often perform better in terms of achieving covariate balance than this simple approach, however. Broadly, propensity score methods can be categorized into matching methods, subclassification methods, and weighting methods (Stuart, 2010). Subclassification methods rely on stratifying analyses within subgroups of individuals with similar propensity scores (Rosenbaum et al., 1984). Weighting methods include inverse-probability-of-treatment weighting (IPTW) used to estimate marginal structural models (MSMs) (Hernán et al., 2000, 2006).

Propensity score methods can (and should) be integrated into other analyses, for example, regression or g-computation (Ho et al., 2007). Combining the propensity score approach with subsequent outcome analysis can be thought of as doubly robust because the propensity score approach models confounding in the treatment model while the outcome-based analysis models confounding in the outcome model. In some situations, we may have a better understanding of the exposure process – and the propensity score model may be more likely to be correctly specified – whereas in other situations we may have a better understanding of the outcome determinants, thus the outcome model may be a better bet. With a doubly robust model, correctly specifying either the exposure or outcome model provides unbiased estimates of the effect of interest, even if the other model is misspecified.

Propensity score weighting methods such as IPTW, which are typically used to estimate MSMs, can be applied in longitudinal studies to address time-varying confounding where the time-varying confounder is affected by the previous exposure and affects subsequent exposure and outcome (Robins et al., 2000). Because of the dynamic and powerful nature of many social exposures, time-varying confounders affected by prior exposure are a common problem and MSM-type approaches are proving promising in longitudinal studies. For example, Cerda et al. reexamined the link between neighborhood poverty and drinking frequency using the longitudinal Coronary Artery Risk Development in Young Adults study, accounting for the potential time-varying confounding variables of income, education, and occupation using MSMs (Cerdá et al., 2010). These types of variables pose particular challenges in neighborhood studies where they can act both as confounders and also mediators. Cerda et al. found that the MSM estimate that accounted for the time-varying nature of these confounders was stronger than the typical regression estimate and statistically significant. Gilsanz et al. implemented MSMs to evaluate the effects of changes in depressive symptoms on risk of subsequent stroke (Gilsanz et al., 2015). MSMs were necessary in this context to address time-varying confounding, in that the resolution of depressive symptoms is influenced by past symptom severity, behavioral patterns, and comorbidities, that are themselves consequences of depression. Although dozens of previous studies demonstrated that individuals with elevated depressive symptoms also had high risk of stroke, these analyses did not directly address the most relevant question for guiding treatment: does reduction in depressive symptoms reduce stroke risk? Gilsanz found that stroke risk remained elevated even for individuals whose depressive symptoms resolved between biennial assessments. Bias from time-varying confounders affected by prior exposure can also be addressed in outcome modeling, through structural nested models using g-computation (Robins, 2000). Although doubly-robust extensions of traditional MSMs exist (Vansteelandt et al., 2014) we have not seen them applied in social epidemiology.

In addition to addressing bias, propensity score methods help avoid violations of what is called the positivity assumption (Pearl, 2009), which requires that people in every stratum of covariate values have a positive probability of being in the exposed group and

Please cite this article in press as: Glymour, M.M., Rudolph, K.E., Causal inference challenges in social epidemiology: Bias, specificity, and imagination, Social Science & Medicine (2016), http://dx.doi.org/10.1016/j.socscimed.2016.07.045