The Relevance of Social Epidemiology in HIV/AIDS and Drug Abuse Research

Richard Rothenberg, MD, MPH
Emory University School of Medicine

“For the fabric of this universe is like a labyrinth to the contemplative mind, where doubtful paths, deceitful imitations of things and their signs, winding and intricate folds and knots of nature everywhere present themselves, and a way must constantly be made through the forests of experience and particular natures, with the aid of the uncertain light of the senses, shining and disappearing by fits.”

Francis Bacon, The Great Instauration of Human Control in the Universe, 1620

Relevance is no longer the issue for social epidemiology.

Implicit in this chapter’s title is that social epidemiology is somehow different from “standard” (or “traditional” or “regular”) epidemiology. Perhaps it required recognition as its early ideas were developed,1–4 but, as Kawachi succinctly summarizes,5 the field is clearly established. Questions of its relevance, like the endless discussions of its definition, are no longer relevant. In fact, epidemiology, in its logistic lethargy, has been reenergized by the issues that “social” epidemiology brings to the fore: methods—the need to deal with different logical types of data in a single framework;6–10 theory—the need for some guiding thoughts that provide testable hypotheses and logical syntheses;11,12 and philosophy—the meaning and role of causality in epidemiologic thinking.13–19 In other words, social epidemiology differs from traditional epidemiology only because it is harder. Consider for example, the factors that conspire to give HIV special status among diseases,20 shackling investigation and intervention, and an equally complex set of factors that contort the issue of drug abuse.

In fact, the labyrinth and intricate folds are never more in evidence than when we try to understand the interacting intricacies of substance misuse and HIV transmission. For the former, Galea and colleagues21 provide a scholarly critique of nearly 100 recent publications that deal with the role of social factors in substance use initiation, cessation, and relapse. By their own admission, their review is not comprehensive and requires a valuation of what to include. (It is hard to imagine a comprehensive review. A quick look at PubMed fetches 15,858 entries for “sexual orientation” and 12,481 for “personality” when either is combined with “HIV/AIDS” and “Drug Abuse”.) The field is already past the point where complexity is readily encompassed.

Correspondence and reprint requests: Richard Rothenberg, MD, MPH, Emory University School of Medicine, 69 Jesse Hill Jr. Drive, SE, Atlanta, GA, 30303, Phone: 404-616-5606; E-mail: rrothen@emory.edu

Publisher’s Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
The issue is dealing with complexity

Taking a somewhat different tack, Poundstone et al.\textsuperscript{22} offer a nested Venn diagram of the social epidemiology of HIV transmission. Their diagram (see the Figure 1 that appears in their paper, not reproduced here) places the set called \textit{HIV incidence} within increasing inclusive sets: \textit{HIV Transmission Dynamics} within \textit{Individual Factors} within \textit{Social Factors} within \textit{Structural Factors}. Each set has a number of elements, and each element, a set of sub-elements. Taken as a list, the 36 sub-elements they enumerate are testimony to the daunting breadth of the field, though they are obviously incomplete (the sub-elements for \textit{Cultural} and \textit{Social Capital} effects are left out—too numerous to mention, perhaps).

The visual model has become an important tool in social epidemiology for dealing with complexity.

Such diagrams have become the stock in trade of social epidemiology. In a thoughtful critique, Kaplan\textsuperscript{11} revisits five graphical models, each with an extractable list of critical social factors. All such graphs are an attempt to reduce multiple dimensions to a more manageable two, often with some success. The world may not need yet another graphical model, but Figure 1 is offered to illustrate several problems with this approach. Figure 1 is not intended as a graphical representation of causality,\textsuperscript{14} since it lacks both an underlying animator and the arrows to provide orientation. (Note that several of those reproduced\textsuperscript{23,24} or generated\textsuperscript{25} by Kaplan are meant to be causal graphs.) Such models imply a mechanistic view of the process, and are contingent on the notions of causality alluded to earlier. As we will discuss, such a mechanistic causal framework may ultimately be suited to only limited aspects of the issue at hand.

Visual models may fall short because they do not capture causality, they are not actually fractal, and they finesse the complexity.

Figure 1 serves as a fractal metaphor, using the artifice of embedded pentagons. Krieger\textsuperscript{26} suggested such a metaphor to highlight the repetitive nature of multiple interactions at multiple levels. Though the symmetry is seductive, such diagrams can be no more than metaphor, since the essence of fractal behavior is recurring regularity. A defined recurring pattern at each hierarchical level would actually be a welcome simplicity, but the true situation is more complex: the five sides are arbitrary and a large and variable number of factors show up as one drills deeper at any point. Though the fractal metaphor has visual value, the true situation is untidy, and suggests that social complexity cannot be easily confined, a point often lost in modeling (or perhaps more fairly put: a point often bypassed in modeling, whose purpose is reasonable simplification in the service of comprehension\textsuperscript{27}).

A case in point would be the behavioral pentagon. It is depicted as a 5-clique (in the parlance of network analysis, a set of five nodes [usually people but in this case, items] with a path between each possible pair). A closer examination of the actual risks involved (Figure 2) suggests that simultaneous multiple risk taking is the rule, rather than the more traditional hierarchical structure imposed on the epidemic during its early years.\textsuperscript{28} Rather than a pentagon, a large polygon of interacting risks is a more likely representation of the reality (not all of which are shown in Figure 2, and not all of which were explored in this study\textsuperscript{29}). This multiplicity has been demonstrated in other studies from the Atlanta area,\textsuperscript{30,31} as well as numerous other examinations of the inner city risk environment (references on request). In the larger context of social epidemiology (Figure 1), these behavioral interactions are actually some of the simpler ones to deal with.
Social epidemiology and social network analysis appear to occupy different orbits.

Finally, Figure 1 highlights the importance of social network and geographic factors in the overall social epidemiology of drug abuse and HIV transmission. Social network analysis, whose earliest use is usually attributed to Moreno\textsuperscript{32} despite antecedents,\textsuperscript{33} is a set of quantitative tools applied to the “web” of human interactions. Though primarily an approach to sociologic questions, it has found application in recent years to such diverse areas as cell biology and antiterrorism efforts. Klovdahl, in 1985, introduced social network methods to the study of HIV/AIDS, by reanalyzing information about an early cluster of 44 men with AIDS.\textsuperscript{34} Subsequent work on many fronts has explored the role of social, sexual, and drug-using networks in the dynamics of disease transmission.

There is a curious disconnect between the worlds of social epidemiology and social network analysis. Though Poundstone et al. include it in their diagram and discussion,\textsuperscript{22} references to social networks from the literature of social epidemiology are rare. For example, in their discussion of causality in social epidemiology, Kaufman and Cooper\textsuperscript{15} state that “…(social epidemiology) has been less productive, however, in the development of a similarly rigorous body of theoretical work on the quantitative and structural analysis of social interactions, which would underlie any analysis accommodating interunit dependencies. Some preliminary work along these lines has appeared within the field of sociology [reference here to Wasserman and Faust’s textbook\textsuperscript{35}], but as yet we know of no application of this type of scholarship to the problems confronting social epidemiology.” (page 120). Possibly, their use of the term “interunit” referred to the social factors that have been called “sub-elements” in this discussion, but if by units they mean people, then they are ignoring a universe at least as populated as their own. (Their reference to the superb standard textbook in network analysis is akin to describing the field of linguistics with a reference to the dictionary, but in fairness, one rarely sees the term “social epidemiology” in the network literature either.)

A greater mutual recognition of social networks and social epidemiology could be of value in dealing with the complexity.

The more important issue is that network approaches can help in dealing with the complexity of factor interaction. The concern with the relationship between “macro” and “micro” events is deeply embedded in sociology and network analysis, just as it is in social epidemiology (as so clearly stated by Kawachi in the editorial that opened the Social Epidemiology section in Social Science and Medicine\textsuperscript{5}). This concern was expressed by Granovetter in 1973,\textsuperscript{36} when he said: “A fundamental weakness of current sociological theory is that it does not relate micro-level interactions to macro-level patterns in any convincing way. … I will argue…that the analysis of processes in interpersonal networks provides the most fruitful micro-macro bridge” (page 1360). Thirty years later, Morris\textsuperscript{37} succinctly hypothesized that the bridge is the relationship between local rules (meaning the choices that individuals make) and the global properties of networks that such choices generate. The social, sexual or drug-using networks that results from the agglomeration of personal choices will determine disease propagation (to which the biology of transmission is only one of the contributors). The interdigitation with social epidemiology is clear: Social factors influence the choices people make and social factors are influenced by the network structures that result from such choices. For example, think of the interplay between the network of inner city dwellers who are at risk for HIV and drug abuse, and the network of influential societal leaders who make laws affecting those inner city dwellers. The result is a recursive process rather than a causal chain. Note as well that in this construct, no spider is required for a web of causality,\textsuperscript{26} analogous to the argument that evolution requires no unseen hand.
Insertion of networks into the overall thinking of social epidemiology could have some important ramifications. First, just as “traditional” epidemiologists have sought new methods (or really, revived and expanded old ones) to deal with social issues, network analysts have had to move beyond standard epidemiologic approaches, and toward the development of alternate approaches to sampling,38–40 hierarchical modeling in a network context,6 and the exploration of alternative models for understanding disease transmission.41–43 In addition, the recent contributions of statistical physicists, whose graph-theoretic orientation has produced new methods and insights,44–49 has expanded the vision of how to deal with complex systems well beyond the borders of traditional epidemiology.

Similarly, the insertion of network and geographic thinking into social epidemiology has value in understanding the propagation of disease. Social epidemiology has an intense concern with geography, but usually in the form of trying to define the role of neighborhood (or some other jurisdictional unit).4,13,19,50–52 Much of this work has focused on gradients (social, economic, or disease-specific), or on the combined influence of individual and social factors within small areas using multilevel modeling techniques. Like social epidemiology, network approaches focus on geography, but largely on the role of geographic contiguity in disease transmission. Zenilman and colleagues measured the distances between persons with gonorrhea and their partners.53 They demonstrated significantly smaller distances between dyads within core areas of Baltimore compared to non-core areas, reinforcing the importance of local neighborhood and geographic compactness in the dynamics of gonorrhea transmission. In an expansion of this concept,54 comparison of social distance (measured as the geodesic, or smallest number of steps from one person to another) and geographic distance (measured as kilometers between residences) showed that persons involved in a high risk network demonstrated \textit{remarkable} contiguity within a well defined neighborhood (Figure 3). The implication of this observation for maintenance of endemic transmission are clear: persons in this milieu have a higher probability of choosing as new partners persons to whom they are \textit{already connected}, and who already share a higher risk. In the larger context of social epidemiology, the factors that foster such geographic proximity are thus part of a definable mechanism for maintaining endemicity.

The notion of network connections between micro- and macro-level social phenomena is associated with some alternatives to current thinking about causality as well. The latter is shaped by traditional views of cause and effect (smoking and cancer), recently modified by a questioning of the fundamental epidemiologic process. As Kaufman and Cooper15 point out, the counterfactuals to many of the important variables in social epidemiology have little meaning. The fundamental epidemiologic goal is to compare a person to himself with one characteristic changed, in order to assess the importance of that characteristic. Since we usually cannot do that, we compare that person to someone who is as like him as possible in all characteristics except the one of interest. This comparison can work for some variables, but not for others (for example, it is impossible to isolate “race” and pretend that a Black person could have been White with nothing else different). For this reason, according to the authors, some causal conclusions are compromised, and social epidemiology is left with an inadequate causal framework.
These thoughts have special cogency in the overall framework for causality in epidemiology. As Greenland has demonstrated, the major methods for assessing causality (graphical, sufficient-component, potential outcome, and structural equation modeling) are interrelated and play a key role in epidemiologic thinking. But, as alluded to earlier, such approaches may be relevant, but not optimal for the purposes of social epidemiology. In discussing causality for complex systems, Wagner\textsuperscript{55} writes:

“In the social sciences, it seems that the linearity of causal models is often due to a lack of insight into functional relations among the variables of a model. This lack of insight probably exists for good reasons, such as the inability to do experiments, and methodological difficulties in collecting even observational data. Given this lack of insight, the most parsimonious assumption about a system is that it is in equilibrium, and that the interactions among its state variables and parameters in equilibrium are linear.”

He and others\textsuperscript{56,57} point out, however, that these complex systems are distinctly nonlinear and can only be approximated by linear causal pathways under narrow conditions. As Philippe and Mansi state:\textsuperscript{57} “Strong nonlinear and chaotic dynamics as well as linear dynamics occur when systems exhibit many, if not all, of the following characteristics: multiplicity of components, of functions, and of initial and contextual conditions; diversity of links among the variables that are measurable at multiple micro/macro levels as well as presence of multi-level links between components across different scales of organization of the system. Feedback loops with sensitivity to minor variation in initial conditions are also present.” The statement sounds remarkably like a description of social epidemiology, whose underlying theme is the complex interrelationship of diverse factors. The essence of this interrelationship is that multiple factors change jointly, that is, that their distributions have an impact on one another that cannot be captured by correlational or other linear analysis. The endpoint of this reasoning is that the “system” of interacting factors is the cause.

This is more than an academic concern for understanding drug use and HIV transmission. Because the direct connection between social factors and a particular individual are often tenuous, multiple, difficult to measure, and perhaps impossible to predict, we never quite know at which interface we can actually use the word cause (here in the sense of “intervention”\textsuperscript{14} or something that can be changed). A set of interfaces, both simultaneous and sequential, would include molecular, biochemical, physiological, interpersonal, social, and societal—all of which might claim primacy. The answer might be all, each, or none, depending on the complex set of causal criteria one adopts.\textsuperscript{16} The implications for promoting structural changes in the face of all these other things are considerable: from the point of view of decision makers, causal ambiguities enable deniability.

Narrative, rigorously connected to theory and data, can be the vehicle for conveying causality.

The intellectual arc from complex systems to convincing causal explanations to resultant structural interventions for drug use and HIV transmission is under construction.\textsuperscript{58} But if the overall system is in fact the cause, one important tool for social epidemiology is narrative—the description of that system. Though not a fundamental solution to the problem of causality in complex, nonlinear systems, the notion of a story is achieving greater credibility, even in mathematical circles,\textsuperscript{59} as a method for providing a path through the labyrinth. In fact, such story telling may be susceptible to a rigorous and quantitative approach. For example, in a creative amalgam of historical scholarship and network techniques, Bearman, Faris, and Moody\textsuperscript{60} reexamined historical accounts from persons who had lived through tumultuous times, and used network methods (the technique of bicomponent analysis) to assess the
robustness of a set of interrelated chronological events as a causal description of the history. Another recent example illustrates both the power of including network methods in causal analysis, and the power of rigorous narrative. Moore and colleagues performed a textual and network analysis of recent major contributions to the assessment of social capital in public health. The network analysis revealed that the communitarian view (social capital as a neighborhood-level variable) dominated thinking, and the importance of network–informed views of social capital was marginalized. They suggest this phenomenon as a cause (in the less formal sense of cogent explanation) for dissent and disappointment with the social capital approach.

In an analogous way, an area for development in social epidemiology would be the network-based linkage of multilevel factors to form a coherent, causal account of the elements that influence the outcomes of interest—here, drug use and HIV transmission. The challenge for the amalgamation of social epidemiology, social network analysis, and causal statistical approaches is the development of a rigorous approach to credible narrative. Such an approach would draw on a number of different sources; would be theoretically-based and empirically-validated; would speak to the larger overall system, rather than chains and links, by including the nonlinear interaction among the parts; and would provide direct connections to policy decisions. Refinement of the techniques that go into such narrative creation may be a fruitful pursuit for social epidemiology which, while preserving its considerable investment in statistics and theory, can move closer to a nonlinear conceptualization. Epidemiologists have nothing to lose but their causal chains.

Acknowledgements

No financial conflict of interest was reported by the authors of this paper.

References


Am J Prev Med. Author manuscript; available in PMC 2007 September 12.


42. Snijders, TAB.; van Duijn, MAJ. Conditional Maximum Likelihood Estimation under various specification of Exponential Random Graph Models. 2002.


56. Fellman, PV.; Wright, R. Institute of Social Psychology. London School of Economics; Modelling Terrorist Networks. 7-18-0005


Figure 1.
An expanded view of the factors relevant to understanding drug abuse and HIV transmission
Figure 2. Interaction of sexual and drug injection risks in a group of inner city persons in Atlanta, Georgia, 1995–1999
Data from Rothenberg et al.29
The numbers are the percentage of persons (men and women) who engage in the specified behavior. Substantial differences exist in the proportion of men who have sex with men (9.4%), who admit to insertive anal intercourse (28.3%) and who have a same-sex sexual orientation (6.9%). The difference results from respondents’ perception of the sexual implications of insertive anal intercourse.
Figure 3. Geographic placement of dyads from a population of persons at risk for HIV transmission on the map of Colorado Springs.

The dense connections are found in a 6 km square area of downtown Colorado Springs. The lines all represent single dyads. If the entire network is placed on the graph, or some manageable subset such as geodesics of 3 or less, an identical pattern of connections is seen.54