Comment: Emerging Objectives and Methods in Epidemiology

Susser and Susser have analyzed epidemiology's past and find this discipline currently in transition from an era employing a "black box" paradigm to an era of "eco-epidemiology" with a new paradigm.[1,2] They admonish us to choose directions for this paradigm that keep a central focus on public health. In that spirit, I elaborate on a new paradigm that is compatible with the Sussers' discussion. I then illustrate its value for the study of infectious diseases.

The Current Transition in Epidemiology

Epidemiology is in transition from a science that identifies risk factors for disease to one that analyzes the systems that generate patterns of disease in populations. The focus of epidemiology is expanding from relationships between exposure and disease variables to the analysis of the systems that give rise to exposures and through which those exposures act to cause disease. Our view of populations is being transformed from a collection of individuals to a set of interactions between individuals—from an additive heap of risk factor effects to a nonlinear system with multiple control mechanisms and leverage points. Dynamic systems models are supplanting fixed mathematical relationships for examining relationships between exposure and disease and predicting the effects of interventions. Issues regarding the conformation of systems are bringing a whole new set of phenomena under the purview of epidemiological investigation. These include contact patterns between individuals as well as the social structures through which individuals affect each other.

This transition is being advanced by the adaptation of methods developed in other disciplines. Eventually it may be further advanced by new developments in the analysis of complex adaptive systems[3] and by addressing questions regarding how disease systems evolve.[4]

A possible future for epidemiology is discerned by considering a similar transition in biology. The identification and classification of species was at one time a dominant activity in biology. The formulation and evaluation of theories involving biological and ecological systems are now much more central to biology. Population biology, evolutionary biology, and ecology have become firmly established traditions as a result of this transition.

In the course of biology's transition, the identification and classification of species was not eliminated as an important activity. Rather, it was transformed. Species identification and classification are now undertaken not only in the older context of descriptive biology, but also within a more comprehensive theoretical framework. Rather than only identifying species they encounter in the field, biologists now explain which species exist on the basis of evolutionary or ecological theory. The examination of ecological and evolutionary relationships plays a central role in the decision as to whether a group of organisms represents a new species.

Likewise, the transition in epidemiology will transform rather than displace the activity of risk factor identification. Better theoretical structures will point our search for risk factors in more productive directions, and decisions on causality will benefit by being put in the context of broader theory. But above all, more opportunities for disease control will be uncovered by addressing a new set of questions regarding the nature and behavior of the systems in which disease arises.

The transition from risk factor detection to systems analysis is equally important for infectious and noninfectious disease epidemiology. The need for it has been made especially clear by the social epidemiologists.[5,6] They point out how the old paradigms of epidemiological inquiry focus too much on factors that are identified by examining individuals, and they provide rough outlines for new paradigms and methods that will provide more inclusive approaches to social epidemiology. Meanwhile, this transition in infectious disease epidemiology has advanced to the point where it can provide experiences and methods of use to other areas of epidemiology. An introduction to the analysis of infection transmission systems is provided by Roy Anderson and Robert May in their book *Infectious Diseases of Humans: Dynamics and Control.*[7]

Systems Analysis in Infectious Disease Epidemiology

Black box–era methods are founded on an assumption that is inconsistent with the transmission of infection, namely that the outcome of exposure in one individual is independent of outcomes in other individuals.[8,9] This inconsistency makes the detection of many risk factors impossible and distorts estimates of the effects of others.[8,9] Why, then, do infectious disease epidemiologists use the black box approach? Two reasons stand out.

First, there are many personal behaviors and environmental contaminations causing infections that can be identified by the black box approach. For a good number of these, mere identification can lead to effective disease control activities. No analyses of how the risk factors act in a larger system and no quantitative predictions of effects from infection control programs are needed. This is the case when the benefits of eliminating a risk factor are clearly greater than the costs or when people will readily change their behavior to avoid a newly identified risk factor. Even though systems analyses methods might detect a greater range of infectious disease risk factors and be of greater value in designing efficient interventions, there are enough easily controlled risk factors that will be identified by black box methods to justify their continued use.

Second, few epidemiologists have acquired the skills needed to analyze transmission systems. Epidemiologists have a knowledge of transmission system elements and a feel for system behavior that most mathematicians engaged in systems analysis lack. Epidemiologists are familiar with the population behavior of infections in a variety of endemic and epidemic situations, and they understand the nature and behavior of infectious agents and of immune responses. But our understanding of transmission systems remains primitive because epidemiologists do not know how to apply their knowledge of system elements to the analysis of the system as a whole. Epidemiologists have not developed their ability to judge which epidemiological observations are most important from a systems point of view and to then abstract those important elements into a mathematical or a computer model on which a system analysis can be based. The skills they need have been expressed by John Holland as follows:

> In building . . . a model, selection is critical. . . . The model (cartoon) can be

Editor's Note. See related articles by Susser and Susser (p 668 and p 674) in this issue.

> more, or less, faithful to the original and, as always, which it is depends on the purpose of the model (cartoon). We may opt for simplicity, or even distorted similarity, at a cost in faithfulness, in order to emphasize some basic element. Newton, in building his models, ignored friction in order to get a more definitive look at momentum. His slightly unfaithful model emphasizes the principle that "bodies in motion persist in that motion, unless perturbed by forces." Aristotle's earlier, more faithful model implicitly included friction, leading him to enunciate the "basic principle" that "all bodies come to rest." Aristotle's model, though closer to everyday observations, clouded studies of the natural world for almost two millennia. Model building is the art of selecting those aspects of a process that are relevant to the question being asked. As with any art, this selection is guided by taste, elegance, and metaphor; it is a matter of induction, rather than deduction. High science depends on this art.[3]

Knowing how to capture the essence of a phenomenon is crucial for a system analysis. The black box–era models that epidemiologists have employed to analyze infectious disease data miss the essence of transmission. In fact they assume away transmission.[8,9] The paradigm of the black box era focuses attention on risk factors that might generate disease in individuals and that thereby create associations between the exposures and diseases of individuals. Consequently this paradigm has led epidemiologists to disregard two crucial elements in infection control: (1) the population pattern of who contacts whom, and (2) risk factors that reside in the infected rather than the susceptible individual.

The problem with ignoring contact patterns can be appreciated by imagining two almost identical populations. For both populations, suppose we have information from each individual on the exposures they experience that affect their risk of infection. These would include environmental contaminations, personal behaviors, physiological states, or anything that could be ascertained by examining either the individuals or their environment. Within populations, each individual might differ from each other individual. But suppose each individual in one population is exposed to exactly the same risk factors as a corresponding individual in the other population. That means that traditional epidemiological data sets from these two populations would be identical. Despite this identity, the two populations can have vastly different levels of infection transmission depending upon who has contact with whom.[10–12] In one population, high-risk individuals may be linked to each other in a core group that sustains transmission, whereas in another population, high-risk persons may not have these links. Black box–era analytic methods, which conceptualize risk as an individual-based phenomenon, cannot capture risk that is determined by how individuals are connected. Transmission system models, on the other hand, express their basic parameters (i.e., contact rates and transmission probabilities) in terms of links between individuals.

The deficiencies of black box–era approaches are especially evident in the measurement of vaccine effects. The paradigm of effects being manifest in the individuals at risk of infection, rather than in the system that circulates infection, has caused some of the most important effects of vaccines to be ignored. Vaccines stimulate immune responses that help to control infection after the host takes up an agent. A vaccine-induced immune response may control some agents before they have any consequences for the individual or for the transmission of infection. On the other hand, it may only decrease the contagiousness of infection without eliminating infection. Such a decrease in contagiousness could provide a crucial element in the control of transmission, especially for HIV infection, where a vaccine that prevents no infection could stop the HIV epidemic merely by reducing contagiousness during primary infection.[12–14]

The importance of primary HIV infection in sustaining the circulation of HIV in a population is missed by the paradigm of the black box era for two reasons. First, exposure to a partner with primary infection is very difficult to ascertain. Second, and more importantly, it is not just the infections directly caused by primary infection that generate its dominance of HIV transmission dynamics. Primary infection can dominate those dynamics even when much more virus is excreted in other stages of infection.[12] That is because transmissions during primary infection are more effectively connected into chains that rapidly disseminate infection. The patterns of connection between individuals underlying those chains are ignored by the black box–era paradigm. A transmission system analysis, on the other hand, makes the dominance of primary infection stand out clearly.[12,15] Estimating transmission probabilities as a function of vaccination status in both the infected and susceptible individual provides a measure of vaccine effects on contagiousness during primary infection that is completely missed by standard methods.[14]

How systems analysis captures important determinants missed by the methods of the black box era is reflected by the data structure for a black box–era analysis and a transmission system analysis. Epidemiological data is traditionally structured to estimate parameters that relate exposure to disease in individuals, such as odds ratios or risk ratios. Dependent and independent variables are placed in columns, and individuals are placed in rows. One problem with this structure is that not all data fall into the class of dependent or independent variables. Some data are relevant mainly to the parameters of a transmission system, namely, contact rates between different classes of individuals and transmission probabilities during contacts.

How long individuals courted before having sex is a case in point. Such a variable might generate a weak association with infection when it is treated as a characteristic of the at-risk individual rather than as a descriptor of the relationship between two individuals. The same variable, in contrast, could play a central role in helping to determine who is having sex with whom in a transmission system analysis. A variable like courtship time should thus be used to reflect how individuals in different rows are connected to each other rather than merely examined for its association with infection.

The social settings in which individuals form sexual partnerships provide another case in point. When treated as an independent variable in a traditional analysis, a social setting might have a negative association with infection at one point in an epidemic and a positive association at another.[8] That is because the number of infected individuals in a particular social setting can vary dramatically over the course of an epidemic. Sex in social settings where oral sex is predominantly practiced might thus appear to be safe in a traditional analysis conducted early in the HIV epidemic.[8] A transmission system analysis, however, might reveal that sex in such settings could later on be crucial in sustaining chains of transmission in a population. The traditional analysis misdirects us away from a crucial control issue because the parameters it estimates, odds ratios for instance, do not reflect the structure and function of a system: they merely reflect transient relationships in the data.

Implications for Epidemiological Training

For this new era of systems analysis to emerge, epidemiologists must acquire skills not taught currently in their training programs. These include (1) facility in abstracting the essence of systems into models; (2) an approach to developing hypotheses about the dynamic systems that generate patterns of disease in populations; (3) an ability to explore the potential behavior of dynamic systems through model analysis or simulation; and (4) a capacity to use model analyses and simulations in the design of field studies. At many universities, epidemiology students might learn these skills in disciplines other than epidemiology. It would be useful to learn them, however, in the context of epidemiological problems.

To this end, epidemiologists should collaborate with biomathematicians, computer modelers, and systems engineers. This collaboration should encompass both scientific investigation and the development of new courses for epidemiologists in training. Recent technological advances should facilitate the roles of epidemiologists in these collaborations. Programming tools for computer modeling of dynamic and probabilistic systems have been developed to the point where such modeling can be quite feasibly undertaken by almost all epidemiologists. Deterministic models can be constructed with a variety of programs like Stella (High Performance Systems, Lyme, NH). Probabilistic models of populations of discrete individuals can be constructed with simulation packages like those under development at the National Micropopulation Simulation Resource at the University of Minnesota (http://www.nmsr.labmed.umn.edu/nmsr/NMSR.html). Instruction in the use of such tools should be an integral part of all doctoral training in epidemiology. □

James S. Koopman
Department of Epidemiology
University of Michigan
Ann Arbor

Acknowledgment

The editorial assistance of Helen Fox improved this paper.

References

1. Susser M, Susser E. Choosing a future for epidemiology: I. eras and paradigms. *Am J Public Health.* 1996;86:668–673.
2. Susser M, Susser E. Choosing a future for epidemiology: II. from black box to Chinese boxes and eco-epidemiology. *Am J Public Health.* 1996;86:674–677.
3. Holland JH. *Hidden Order.* New York, NY: Helix Books, Addison-Wesley; 1995.
4. Nesse RM, Williams GC. *Why We Get Sick.* New York, NY: Times Books, Random House; 1994.
5. Krieger N. Epidemiology and the web of causation: has anyone seen the spider? *Soc Sci Med.* 1994;39:887–903.
6. Wing S. Limits of epidemiology. In: Wesley RC, Sidel VW, eds. *Medicine and Global Survival.* Vol. 1. London, England: BMJ Publishing; 1994:74–86.
7. Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control.* Oxford, England: Oxford University Press; 1992.
8. Koopman JS, Longini IM, Jacquez JA, et al. Assessing risk factors for transmission of infection. *Am J Epidemiol.* 1991;133:1199–1209.
9. Koopman JS, Longini IM. The ecological effects of individual exposures and nonlinear disease dynamics in populations. *Am J Public Health.* 1994;84:836–842.
10. Jacquez JA, Simon CP, Koopman JS, Sattenspiel L, Perry T. Modeling and analyzing HIV transmission: the effect of contact patterns. *Math Biosci.* 1989;92:119–199.
11. Koopman JS, Simon CP, Jacquez JA, et al. Sexual partner selectiveness effects on homosexual HIV transmission dynamics. *J Acquir Immune Defic Syndr.* 1988;1:486–504.
12. Koopman JS, Jacquez JA, Simon CP, et al. The role of primary HIV infection in the spread of HIV through populations. *J Acquir Immune Defic Syndr.* In press.
13. Koopman JS, Simon CP, Jacquez JA. Assessing contagiousness effects of vaccines and risk factors for transmission. In: Kaplan E, Brandeau M, eds. *Modeling the AIDS Epidemic.* Philadelphia, Pa: Raven Press; 1994.
14. Koopman JS, Little RJ. Assessing HIV vaccine effects. *Am J Epidemiol.* 1995;142:1113–1119.
15. Jacquez JA, Koopman JS, Simon CP, Longini IM. The role of primary infection in epidemics of HIV infection in gay cohorts. *J Acquir Immune Defic Syndr.* 1994;7:1169–1184.