Annu. Rev. Public Health. 1999. 20:15–33

A FUTURE FOR EPIDEMIOLOGY?

S. Schwartz,[1, 2] *E. Susser,*[1, 2, 3] *and M. Susser* [1, 4]
[1]Joseph L. Mailman School of Public Health (Epidemiology); [2]CUES, New York Academy of Medicine, New York, New York 10029, [3]Department of Psychiatry and [4]Sergievsky Center, Columbia University, New York, New York 10032; e-mail (Schwartz): SBS5@columbia.edu; (M Susser): MWS2@columbia.edu; (E Susser): ESS8@columbia.edu

KEY WORDS: methods, paradigms, history, critiques of epidemiology, eco-epidemiology, multilevel epidemiology

ABSTRACT

This paper considers ways of thinking about causes and prevention that could guide epidemiology beyond the present era. Discontent with modern epidemiology, in the face of its substantial achievements, is taken as a sign that the guiding principles of the discipline warrant reconsideration. To begin this task, current practices are placed within an historical perspective, in a review of the dominant ideas of successive eras in epidemiology. Then the premises and constraints of the present era of chronic disease epidemiology, with its risk factor paradigm, are specified. Finally, elements of a causal paradigm for an emerging era are proposed. This paradigm encourages thinking about causes at multiple levels of organization and within the historical context of both societies and individuals. The proposed approach aims to preserve and build on the contributions of past eras, as well as the present one.

INTRODUCTION

This review considers the current state of epidemiology, what its future holds and what epidemiologists might do to meet that future. Although the approach of the millennium provides a symbolic stimulus for self-examination in many endeavors, for epidemiology the timing is propitious. The present era of chronic disease epidemiology has come of age over the past 25 years. The signs are seen in greater consensus and clarity about study designs and such central ideas as causation, confounding, and interaction (53). Multiple texts of different levels of sophistication, where few existed before, reflect this maturation (e.g. 11, 23, 55).

0163-7525/99/0510-0015$08.00

Yet a rising tide of self-criticism is companion to this substantial development. From both within and outside the discipline, a perception grows that epidemiologic research produces small and inconsistent risk ratios, with unclear results that are not easily distilled into cogent public health messages (43). Doubts are voiced about the commitment of epidemiologists to addressing the public health issues of our time and even about the importance of the contribution epidemiology can make to these issues (75).[1]

How are we to understand this ironic mix of intellectual advance and professional dissatisfaction? What does it tell us about the strengths and limitations of current epidemiologic practice? Can we shape the future of epidemiology to enhance the strengths and overcome the limitations?

To address these questions, we first take note of prevailing criticisms of chronic disease epidemiology. We then set current practices against the evolution of the field. Our historical sketch places at its center Kuhn's (30) notion of scientific paradigms that dominate successive eras. This framework gives each period an intellectual history that can be discussed in terms of the questions deemed important, the methods used to study them, and their impact on public health practice. Such a review should bring to the surface some of the assumptions and constraints of chronic disease epidemiology and its risk factor paradigm. Finally, we urge consideration of questions and approaches that may help shape the paradigm of an emergent era.[2]

CRITICISMS OF CHRONIC DISEASE EPIDEMIOLOGY

The sources of discontent regarding chronic disease epidemiology pertain to both the production of epidemiologic knowledge and the purview of epidemiology.

Production of Knowledge

In regard to knowledge production, criticisms of epidemiologic research often focus on two related issues—the problem of detecting small effects and inattention to hypothesis testing.

SMALL-EFFECT DETECTION Critiques of epidemiologic research that focus on the problem of detecting small effects hold that the real triumphs of chronic disease epidemiology, such as those with several cancers and cardiovascular diseases, are past; in the future, relatively small effects will predominate. This is

[1]These critiques have met with vigorous responses that will not be discussed here (e.g. 10, 58).

[2]This paper extends previous work on this topic. While new material has been added throughout, the sections "Premises for a New Paradigm" and "Implications for Epidemiologic Concepts" have been developed specifically for this paper.

problematic because small effects are particularly vulnerable to indeterminacy from confounding and bias (72).

Some critiques from this perspective emphasize the need for methodological rigor and replication. This would help avoid premature announcements of associations between risk factors and disease that so often confuse and alarm the public. In this view, the assumptions of modern epidemiology are sound. The difficulty resides in the reality that potentially important effects are indeed small and difficult to uncover (49) and that strict epidemiologic standards are not always applied but should be (15, 72).

HYPOTHESIS TESTING Another source of dissatisfaction concerns the "black box" of epidemiology in the chronic disease era (62–64). The black box metaphor connotes that risk factor–disease associations are afforded priority over the linkages between them (62, 79). From one perspective, Buck (6) argued for more conscious development and testing of falsifiable hypotheses. This rediscovery of Popperian principles led to a decade of debate in the field about the utility of inductive versus deductive reasoning and the role of verification and falsification in causal inference (18a, 54). Others argued instead for the careful specification of biological pathways (17).

The Purview of Epidemiology

A more radical and increasingly frequent critique questions the adequacy of epidemiology today as a knowledge base for public health. These critics acknowledge that epidemiology has produced information essential for understanding disease etiology and decreasing the burden of disease. At the same time, they hold that many important public health issues are left unaddressed (34, 61, 75) and the potential role of communities has been inadequately considered (80).

A central charge is the general neglect of the social environment in which disease occurs. It is argued that epidemiologists conceive risk for disease as residing largely within individuals and their personal behavior. The interactions among individuals and the interplay between individuals and the environment fall outside the scope and the grasp of most research. Inattention to context leads to a limited and precarious knowledge base for public health action (1, 28, 35, 45, 64, 80).

HISTORICAL OVERVIEW

Better to profit from these critiques, we sketch past paradigms from the 19th century on, placing the current era with its risk factor paradigm in historical context. We begin with a discussion of the utility of a paradigmatic approach.

Scientific Paradigms

Science does not merely describe nature as it exists. Rather it organizes the limitless observations that can be made toward the ends of understanding, prediction, and control. Scientists are therefore trained in "directed and restricted perception" (16). The dominant conceptions of each era—its paradigm—provide an implicit framework for such perceptions and a coherent structure for scientific thought. At the same time, a paradigm sets unspoken limits on the questions, methods, and concepts that are deemed legitimate. Anomalous observations are often overlooked, underplayed, or reinterpreted to fit.

Dominant paradigms therefore generate a tension between the values of innovation and tradition. Indeed, innovation, prized and resisted at the same time, has been a frequent theme in the sociology of science (2, 85). For epidemiology, an archetypal case is the long controversy about smoking as a cause of lung cancer (65a), most recently described in papers by Doll (12a) and Wynder (83), pioneers in its discovery.

Past Paradigms

Whether a reigning paradigm is overthrown in revolution (30) or eroded by attrition, the sequence of paradigms in successive eras illuminates the evolution of a scientific discipline.[3] In the history of modern epidemiology through the 19th and 20th centuries, three main eras, each with its paradigm, can be discerned: sanitary statistics, infectious disease epidemiology, and chronic disease epidemiology (69).

The miasma paradigm dominated the era of sanitary statistics for most of the 19th century. This theory attributed the diseases of urbanizing societies to foul emanations from contaminated soil, water, and air. Control of disease came to mean above all improving sanitation and drainage.

Toward the end of the 19th century, the revolutionary discovery that microbes cause disease displaced miasma. The germ theory paradigm opened the new era of infectious disease epidemiology. Epidemiologists typically sought to relate a single agent to a specific disease by applying Koch's postulates, as exemplified in his classic paper on tuberculosis (13). To control infectious disease, epidemiologists traced the specific agents and intervened to interrupt transmission.

In the years after World War II, infectious disease epidemiology in turn yielded to chronic disease epidemiology with its risk factor paradigm. In developed countries, familiar infectious diseases—such as typhoid, tuberculosis, diphtheria, and infant diarrhea—had receded. At the same time, chronic

[3]We cannot agree with Kuhn that paradigms are supplanted only by total revolution and, hence, that scientific knowledge is not cumulative from one paradigmatic era to another (79a, 68b).

diseases—such as cardiovascular disease, cancer, and peptic ulcer—were rising at alarming rates. Over a relatively short period beginning in the early 1940s, epidemiologists refocused their attention on the discovery of risk factors for chronic diseases. The dominant paradigm of this era, still with us, is a theory of multiple causes, which eventually crystallized as the risk factor paradigm.[4] Under this paradigm, studies that relate exposures to disease outcomes inform public health interventions to reduce individual risk for disease.

Paradigmatic Constraints: the Example of Peptic Ulcers

As noted earlier, paradigms limit the range of legitimate investigation at the same time as they give it direction. Peptic ulcer illustrates the influence of the dominant paradigm on the realm of inquiry and on causal concepts across successive eras in epidemiology. We begin in the heyday of infectious disease epidemiology and specific-cause germ theory, in the early part of the century, and move on to the era of chronic disease epidemiology and nonspecific-cause theory.

In the early part of the 20th century, epidemiology was defined as the study of infectious diseases. In the words of Wade Hampton Frost, "... epidemiology may be defined as the science of the mass-phenomena of infectious diseases ... concerned not merely with describing the distribution of disease, but equally or more with fitting it into a consistent philosophy" (39). At this time, the number of peptic ulcer deaths was increasing. However, scientific observations were largely confined to physiology, medicine, and the rising specialties of gastroenterology and surgery. In an era focused on infectious disease, peptic ulcer was a typical example of an epidemic disease that appeared to be noninfectious and hence outside the scope of the discipline of epidemiology.[5]

In the interwar period, in the United Kingdom especially, Major Greenwood, John Ryle, and other pioneers of chronic disease epidemiology began to argue for a focus broader than infection. Ryle, in fact, argued from the example of peptic ulcer. After noting the increasing incidence of duodenal ulcer in England, he described it as "a disease which tends to relapse or chronicity and which has no basis in contagion" (57). Later he speculated that "the increasing stress and pace of life in our industrial and mechanical age provide a set of predisposing causes" (57).

[4]In a previous paper (69a), two of us (MS and ES) termed the paradigm of the chronic disease era the "black box." To avoid pejorative connotation, we prefer the term "risk factor paradigm."

[5]Admittedly, this period also gave us the classic work of Goldberger (74). With his social scientist ally, Sydenstricker, he showed that the causes of at least one epidemic chronic disease in the United States, pellagra, were noninfectious. Their imaginative studies demonstrated both the proximate nutritional cause and its origin in antecedent social and economic conditions. But Goldberger rowed against the epidemiologic current. At that time, most investigators in the United States were still seeking to discover the suspected infectious agent.

With the chronic disease era of epidemiology entrenched after World War II, peptic ulcer came squarely within the scope of epidemiology. Studies by Doll and others began to appear. The search for causes was guided by the emerging risk factor paradigm. Lifestyle factors pertaining to individuals were investigated as candidate causes (8, 82a). Despite notable exceptions (71), the impact of broader societal change, which intrigued Ryle, was rarely the focus of direct investigation. Few thought to consider infection.

Ironically, an infectious agent (*Helicobacter pylori*) has recently been established as an important cause of the disease. In retrospect, unrecognized clues to an infectious agent in peptic ulcer resided in both population data and laboratory studies long before (5, 66). The puzzling findings, evaluated in terms of the dominant risk factor paradigm, were overlooked or reinterpreted.[6]

Marshall & Warren (38) reopened the issue. They showed that peptic ulcer was associated with the newly detected *Helicobacter*. These subversive findings at first provoked skepticism and controversy. To dramatize his case in the spirit of Koch's postulates, Marshall went so far as to swallow a glass of solution infected with the bacteria. Acute gastritis followed (37).

Ultimately, the distinction between infectious disease and chronic disease may evaporate for peptic ulcer; it may be both. A compelling hypothesis is that infection with *Helicobacter* early in life carries a lower risk for duodenal ulcers than infection at a later age. Until recently, *Helicobacter* may have been ubiquitous in human populations, with early infection the rule (5). With industrialization and improving standards of living, the average age of initial infection increased, resulting in higher rates of peptic ulcer in the affected cohorts. At later stages of economic development, a large part of the population was not infected at all; rates of peptic ulcer then began to decline.

Implications of Epidemiologic Paradigms for Public Health

In each era, the dominant epidemiologic paradigm has had crucial implications for public health practice. In entertaining a new public health intervention, the strength of inference about causality is a prime consideration. The transition from one epidemiologic paradigm to another is accompanied by a shift in what qualifies as rational public health practice. Such shifts often, if not invariably,

[6]For example, in the early 1960s, population studies demonstrated birth cohort patterns that challenged the notion of peptic ulcer as a disease of civilization (66, 71). The sharply rising trends in mortality and morbidity of birth cohorts born at the turn of the 19th century and attaining middle age in the mid-20th century reached a plateau. They then began to decline in younger cohorts, just as sharply as they had risen. At the same time, although bacteria had been seen and reported in gastric specimens from ulcer patients, the medical community continued to believe that bacteria could not survive in the acid environment of the stomach.

stoke controversy over public health practice across and even within generations. The transition from miasma to germ theory illustrates the point.

The Sanitary Movement gathered momentum in England early in the 1800s and reached a high plateau in mid century (14, 31). The idea of foul emanations from organic matter in water, soil, and air took firm hold among leaders for sanitary reform. They hoped to control the rampant epidemics of an urbanizing and industrializing society by clearing their miasmic breeding grounds in accumulated waste, cesspools, and contaminated water. Edwin Chadwick devised closed circuits for sewage and water supplies. Florence Nightingale endorsed the sanitary program and insisted on fresh air and cleanliness in the care of the sick. William Farr built a national vital statistics system. John Simon created a public health administration.

In counterpoint, beginning at the zenith of the Sanitary Movement's success, the theoretical foundations for the revolutionary shift from miasma to microorganisms were laid in Europe by Henle, Pasteur, Snow, Panum, Koch, Klebs, and others. As Kuhn's theory of scientific revolution predicts, the advent of the new theory elicited conflict between the old and the new. Chadwick dismissed "imaginations such as the theory of germs and spores," and, similarly, Nightingale placed them "on the same footing as witchcraft" (31). Both held firm on miasma until their deaths around the turn of the century.[7]

The germ theory paradigm was first realized in public health practice in the United States, where sanitary reform had lagged well behind the movement in England. In 1892, Herman Biggs became the first director of the Division of Pathology, Bacteriology, and Disinfection of the Department of Health in New York (50, 81). This historic step heralded a new era of practice. Its explicit purpose was to apply the laboratory science of bacteriology to public health. By the end of the century, the shift to a radically different public health based in germ theory, led by Biggs and Park in New York and Chapin in Rhode Island, was well afoot (82).[8] They fought for health departments to reorient themselves

[7]Others in the movement, however, were more eclectic. For instance, John Simon (Chadwick's successor as the chief official architect of public health reforms in Victorian England) was tolerant of competing theories, including germ theory. On the public health front, he argued against a too-narrow focus on drainage and ventilation and for attention to the housing, nutrition, and social conditions of the poor. He never abandoned a broad focus on urban living and working conditions in the face of the triumph of the germ theory and its narrowed focus on the control of specific infections. Chadwick and Nightingale nevertheless condemned his departures from sanitary orthodoxy (14, 31).

[8]Biggs and Park began this work with an outstanding investigation of diphtheria, the first to demonstrate the use of bacterial cultures as a guide to public health practice in the control of an epidemic. They were among the first to document the significance for epidemic spread of silent carriers of infection. Infectious disease control in New York was later emulated city by city in the United States and elsewhere.

with germ theory as their guide to epidemic control, sometimes with unintended consequences (36b). The general sanitary interventions guided by miasma were to be supplanted by specific measures against infectious diseases.

The proponents of this new era launched strident attacks on the old school. Charles Chapin proclaimed in 1902, "The English, who carried the notion of the danger of filth to the extreme, were assumed to be the leaders in public health work .... It will make no demonstrable difference in a city's mortality whether the streets are clean or not, whether the garbage is removed promptly or allowed to accumulate, or whether it has a plumbing law ... some of the most recent works on sanitation still reiterate the timeworn phrase about dirt and disease ..." (59).

The infectious disease paradigm dominated laboratory research in epidemiology well before it could be widely applied in public health policy and practice. To begin with, opposition from academic medicine and public health was strong, and the old school was not without counterarguments. From a population perspective, the facts left room for miasmatic or similarly broad theories. Thus, the spread of infection in certain well-described epidemics could not be traced to contact with an infected individual. The role of silent carriers and the spectrum of affliction with infection had to be appreciated before the puzzle could be solved in theory and in practice (82).

Meanwhile, miasmatists could claim credit for major public health achievements. Today, indeed, many might endorse the general public health measures advocated by the sanitary reformers. Their vision of poor urban conditions generating manifold health disorders is not without merit in modern public health. Once the battles of transition subsided, Chapin himself broadened his perspective and sowed the seeds for the schema of agent, host, and environment (59).

CURRENT PARADIGM

For contemporaries, the features of a reigning paradigm can be difficult to discern. Training in the norms and practices that reflect the paradigm make its premises and constraints seem natural. They therefore remain unquestioned and often unarticulated. Nonetheless, we will try to make them explicit for the risk factor paradigm and thus, we hope, come closer to seeing its limitations and how they might be met.

Premises

The "web of causation" connects multiple causes to a given disease (36a) and is an apt metaphor for causal inference in the risk factor paradigm. In rejecting the notion of single necessary and sufficient causes, this paradigm implies the existence of unnecessary and insufficient causes. This revised concept of cause, together with the associated terminology of "risk factors" to denote causes,

slowly took hold in epidemiology after World War II (68). New research designs and ultimately new analytic methods were required to meet its demands.

A multiplicity of causes signifies, first, that not all of the exposed will get the related disease and, second, that some of the unexposed will get the disease. Once accepted, this premise imposes a logical obligation to compare the occurrence of disorders in persons with and without the risk factor of interest. Thus, the comparative method and the fourfold table of independent and dependent variables are at the core of epidemiologic practice under the risk factor paradigm (42, 65a). The object is to identify specific risk factors amidst a morass of many. In this endeavor, bias and confounding, twin problems that could make comparisons indeterminate, become critical.

The evaluation and testing of integrated theories of disease etiology play a secondary role under this paradigm. Through close attention to study design and analytic techniques, the effects of risk factors are isolated and assessed. Risk factors might be evaluated for causal plausibility—using a set of causal criteria, for instance (21, 27, 68)—but the elaboration of a causal model need not be the motivating force behind the study (79).

Guided by these premises, the methodological achievements of modern epidemiology have been noteworthy. The principles and conditions for valid case-control and cohort designs were clarified (55). Measures of effect commonly used, especially the odds ratio, became better understood (18), as did causal models and criteria that might apply to causal inference and multivariate relationships (52). Sophisticated statistical methods facilitated by computer technology encouraged complex analysis and precision in assessing the effect of particular risk factors.

Constraints

Although the risk factor paradigm has enabled epidemiologists to master the detection of a single risk factor nestled in a background of multiple risk factors, it has also imposed significant constraints. One is a disproportionate concentration on the description of risk factor/disease relationships rather than the explanation of causal processes. Another is a preoccupation with the individual level of organization to the exclusion of other levels.

DESCRIPTION RATHER THAN EXPLANATION Under the risk factor paradigm, the ideal is to create a state of all things being equal (ceteris paribus) between exposed and unexposed populations. This staple of design and analysis manipulates study conditions to isolate the causal factors from those that are potentially confounding or irrelevant. At the same time, however, these simplified conditions obscure the antecedents of the risk factors under study. They also largely neglect the mediators linking the risk factors with the disease.

Indeed, the controlled clinical trial—the best approximation of ceteris paribus and the model toward which other study designs were to strive (68)—is designed for the identification rather than the explanation of causal factors. The sterilized conditions imposed by clinical trials maximize the isolation of specified relationships. The explanation for these relationships, however, resides in linkages along a causal chain, whose detection is hindered by the very structure of this design (12).

Consonant with this approach, modern epidemiology emphasizes confounding and the separation of mixed effects at the unforeseen and unintended cost of de-emphasizing mediation and the linking of effects. Textbooks in epidemiology typically devote large sections to the definition, detection, and analysis of confounding and give little attention to mediation and causal chains.

As concepts of cause have advanced under the risk factor paradigm, the focus has been further narrowed. For example, Rothman's (52) heuristic device of component and sufficient causes has been enlightening regarding effect estimates, confounding, and interaction. Yet this causal model gives centrality to the "set of minimal conditions that inevitably produce disease" (55). By definition, the set of minimal conditions cannot include mediators and antecedents. It is unclear whether this model can accommodate a series of factors in a causal sequence without losing its heuristic clarity. For example, consider a genetic factor that contributes to a behavior that is in turn a risk factor for disease. Once the behavior is included as a component cause, the genetic factor must be excluded as redundant and not in the set of minimal conditions.

SINGLE LEVEL OF ORGANIZATION The second problem of restricted focus, related to the first, is the almost exclusive legitimacy conferred on the study of the individual level of organization. In causal theories, research designs, and analytic approaches, individuals are the preferred unit of interest (27, 36, 68a). Under the current paradigm, questions about macrolevel social and physical environments or microlevel mediators and antecedents are difficult (although not impossible) to frame.

This focus on the individual level is exemplified in the ways that differences across populations are handled. When a population is identified as having a high rate of disease, studies are frequently launched within the high-risk population to detect the responsible factors. For example, in the United States, where obesity rates are relatively high, eating and exercise habits of individuals within the United States have been the locus of research attention. However, to explain a high US rate of obesity, eating habits and activity levels must differ between the United States and the countries used for comparison. Such population differences will not be readily detected in individual-level studies; they are likely to be rooted in social and economic factors that are not suited to investigations using standard designs of chronic disease epidemiology (73).

TOWARD A NEW PARADIGM: MULTIPLE LEVELS AND THE PASSAGE OF TIME

From the articulation of discontents and constraints, the outlines of a new paradigm for epidemiology have begun to emerge. Several recent proposals hold in common a desire to broaden the scope of inquiry in two dimensions especially. First, the search for causes would pursue multiple levels of organization, extending study both to macrolevels beyond (3, 28, 40, 45, 64, 80) and to microlevels within the individual (40, 46). Second, it would elaborate the time dimension to emphasize both histories of individual development and the history of society. Ideally, at all levels accessible to research, consideration would be given to the dynamic processes linking antecedent events and development with later outcomes.

In proposing a new paradigm, we urge that certain questions and ways of thinking be given greater centrality and legitimation. This work does not require us to undermine or abandon advances of the current paradigm. Although the aim is to break out of its limiting assumptions, we will need to preserve its bedrock of designs and analytic techniques. We can and should hold the ground for which current epidemiology is well suited as we move on to new ground and new questions that elude the current paradigm. In doing so, we should recognize that reductionism is an essential research tool but does not serve as an organizing philosophy (32, 33, 70). To examine isolated features of individuals in today's mode is undeniably useful in establishing risks and causes of disease. That step is merely the beginning and not the end of the new task. Commitment to individual-level designs and analyses should not be allowed to obscure the context in which its components unfold and act.

As will be seen, however, the paradigm shift we propose may lead epidemiology beyond what some consider its appropriate purview. Choosing among paradigms is a value-laden endeavor that specifies what is appropriate. In defining legitimate questions and methods, priority is assigned to some aspects of health and disease over others. These priorities, in turn, are based on assumptions and judgments about the class of factors that are amenable and reasonable to change. The political consequences and sources of these assumptions in current and emerging paradigms can be usefully discussed and examined.

Eco-epidemiology

Two of us (MS and ES) have proposed a paradigm along these lines under the name of eco-epidemiology[9] (70). This paradigm addresses the interdependence of individuals and their connection with the biological, physical,

[9]Some authors (76) have misread eco-epidemiology as a call for more ecologic studies. In response, we have underscored that the paradigm calls for examining multiple levels of organization both beyond and within the individual.

social, and historical contexts in which they live. To do so, it encompasses the changeable contributions and effects on the individual level of both macrolevels and microlevels of organization. Firmly rooted in the concerns of public health, the aim is the study of multiple relationships across levels that would contribute to the expansion of our understanding of disease processes.

Eco-epidemiology thus contends that fruitful theories of disease causation and pathogenesis can, in principle, be conceptualized at all levels of organization. Since detectable causes differ across levels, theories at different levels may each point to distinct understandings of disease and prevention (20, 49, 65). The emphasis on the time dimension implies that health and disease in fact involve processes and should be conceived and studied as such. One would aim to assess the reciprocal interpenetration of factors at different levels of organization, over both the life course of the individual and the history of populations (32, 33, 65a).

How to realize such a paradigm shift is not yet clear. As was initially true for each successive era in epidemiology, substantial methodological and inferential barriers need to be overcome. Available research designs and analytic techniques are not well suited to elucidating processes at multiple levels of organization. For the moment, we defer this discussion. Of more immediate concern are the changes in ways of thinking required (19).

Premises for a New Paradigm

In what follows, we propose three elementary premises for a causal paradigm pertaining to levels of organization, the interpenetration of levels, and historical evolution. These will of course need further development as the field moves toward a new approach. We then note some implications of these premises for such basic epidemiologic ideas as risk, disease, and "third" variables.

LEVELS OF ORGANIZATION Causes of disease occur at all levels of organization and not only at the individual level (12b, 20, 49, 60a, 68a, 80). Each level warrants examination for its impact on health. The decision as to which levels to include should be based on the question at hand, the particular nature of the disease, and the pattern of disease rates.

That causes of disease can be distinct at different levels can be understood through the concept of emergent group properties. That is, at each ascending level of organization, unique characteristics confined to that level emerge. To take a simple physical example, water can be chemically defined at the molecular level as a particular constellation of hydrogen and oxygen molecules. The liquidity of water, however, is a characteristic that neither applies to nor is described by the physical assembly of molecules (78).

Similarly, human groups have characteristics that individuals do not. Nations, but not the individuals who populate them, have political structures. Individuals too have characteristics, such as personality and intelligence, that their cells do not.

Even when group characteristics appear to have individual-level analogues, the group characteristic need not be a derivative of the individual characteristics. Both nations and individuals can be considered "aggressive." But this does not mean that militarily aggressive nations are composed of aggressive individuals (47, 84). In other cases, the group characteristic may be derived from the individual characteristic but take on a different meaning. For example, the average income in a neighborhood may determine frequency of garbage pick-ups and the placement of toxic waste. These factors impact individuals living in the neighborhood regardless of their income.

INTERPENETRATION OF LEVELS At the same time that characteristics of each level of organization may be distinct, these levels interrelate in ways that can mutually influence the play of causal factors at each level (32, 33).

The mutual interplay between levels is exemplified in models of infectious disease epidemics. Ronald Ross' (51) theory of "dependent happenings" illuminates the problem of a straightforward sequential causal model in reflecting and predicting patterns of infectious diseases (19a, 26, 68). The risk of infection for an individual is connected with the prevalence of that infection in the groups that surround the individual. The absence of the disease in the contacts of an infected person allows the further transmission of the disease. Herd immunity is a facet of this intimate reciprocal relationship between infection in individuals and groups and sets a prevalence threshold beyond which epidemic spread in a population is blocked.

The phenomenon of dependent happenings deserves fuller consideration in the study of noninfectious diseases, other health-related outcomes, and risk factors. For instance, people who use drugs may recruit others to drug use, thus transmitting this outcome. The overall prevalence of drug use in a community also influences individuals' risk, as when drug exchanges become an important part of the local economy.

This interplay between levels can also be seen in the relation between the behavior of individuals and the characteristics of the community in which they live. For example, psychological factors might induce a propensity for alcoholism, but expression of the propensity depends on the availability of alcohol and the norms surrounding its use. In turn, individual propensities will influence the availability of alcohol in the environment and the norms of drinking (4).

A more complex example is the relation of genetic and environmental determinants. The risk factor paradigm is being extended to include both genetic and

environmental factors, as the use of association studies becomes widespread in genetic epidemiology (25, 44, 48). Nonetheless, the implications of their dynamic interplay—at multiple levels—has yet to be adequately confronted.

First, environments influence the expression of genes, and genetic dispositions influence exposure to environments. But, in addition, effects of genes as well as environments are occurring on multiple levels of organization. A genetic polymorphism may limit the ability of an individual to detoxify carcinogenic aromatic amines. However, only in an environment where this carcinogen is introduced by human economies does it become a potent risk factor for cancer (46). At a different level, the hormonal environment of the fetus may modify the expression of genetic dispositions in sexual development (41). Thus, multiple levels of environments may alter the expression of genetic effects and vice versa. As we gain more control over genetic effects (as with genetic engineering or cloning), these processes will become even more inseparable.

HISTORY History, societal and individual, deserves a central place in the search for causes. In societies, patterns of disease change over time and, in a broad sense, their causes are historically contingent. In individuals, too, states of health and their antecedents are historically contingent. At any given time, patterns of disease and health states are the result of dynamic antecedent processes (20).

With regard to the history of populations, studies that cross historical periods are needed to examine the effects of that history. Aspects of disease development that are otherwise hidden could then be illuminated. Under the risk factor paradigm, we most commonly—and conveniently—study the development of disease in individuals over a period short enough to regard historical time as constant. Effects of causal factors that distinguish one historical period from another are likely to be obscured.

In a typical study under the current paradigm, for instance a study of hypertension, the focus is on factors differentiating individuals at one moment in time, such as salt intake or genetic factors. In studies from a historical perspective, the focus is more likely to be on such historical changes as urbanization and industrialization that differentiate time periods and societies and could underlie changing patterns (60). It is not vain to consider the implications for intervention. Rather than considering the features of such historical phenomena as fixed, it may be possible to modify them so as to lessen the adverse impact on health.

With regard to individuals as well, health states change over time and are contingent, so that we need designs that can trace the emergence of health states over the life course from conception on. An individual's history and interactions with the environment leave somatic and psychic imprints. Each risk factor has antecedents, and, in the personal history of the individual, these

antecedents form the context in which more proximal factors act. Merely to relate risk factors to later disease is not sufficient for this purpose. Such crucial aspects as the impact of developmental stage and the chain of causation fall out of consideration. Investigations need to elicit early life exposures in individual development and locate their role, if any, in the causal chain for diseases with adult onset (29).

Implications for Epidemiologic Concepts

As we move toward a new paradigm, the building blocks of our methods need to be compatible. Although we must surely preserve what we know, some of our present formulations of risk, disease, and "third" variables pose barriers to thinking about multiple levels of organization and process.

RISK Current formulations make the simplifying assumption—for some purposes very useful—that individuals have a stable underlying risk. In addition, cumulative incidence, the proportion of a fixed population who become diseased, is typically conceptualized as expressing the average risk of individuals rather than influencing it (55).

To an important degree, however, an individual's risk of disease is an ever-changing probability based on interactions among the individual's personal history, biology, and physical and social environments. To conceptualize individual risk as in flux is a step toward thinking about process and historical contingencies. To think of an individual's risk as context dependent encourages recognition that group characteristics can influence the disease experiences of individuals.

DISEASE The definition of disease always reflects, to some degree, the dominant causal paradigm. The transition from miasma to germ theory was followed by a redefinition of many common diseases. Today, as genes become more accessible, many diseases are being redefined in terms of their genetic component. For example, a spectrum of schizophrenia-like disorders appears to aggregate in the families of individuals with schizophrenia (24). It is argued that genetic factors in schizophrenia may be more readily detected by using this broad definition and even that the disorder could be redefined in this way.

Full consideration of multiple levels calls for flexibility in defining outcomes. The health outcomes used in causal research should allow for examination of impacts at different levels of organization. To appreciate the full impact of risks at all levels of organization, the study of clinically defined disease entities must be augmented with the study of health outcomes defined in other ways, from physiological systems to social roles.

An intermediate step, urged by Cassel (7) and Stallones (64), would be to proceed more frequently from the theoretical action of the risk factor to

examine multiple disease outcomes (e.g. smoking-related diseases). With this approach, the full impact of social and economic factors, often hidden behind factors more proximate in the causal chain, can emerge. In Cassel's example, the role of socially induced stress in the high rates of infectious disease among recruits exposed to the boot camp experience was more readily apparent when the outcome examined was a range of infectious diseases rather than a single one (9).

"THIRD" VARIABLES The hierarchy among "third" variables, those associated with both the risk factor and disease of interest, needs recasting. The disproportionate emphasis on confounding over mediation ascribes greater importance to the separation of causes one from another than the understanding of the relationships among causes and the processes that link them. Mediation deserves as much attention as confounding, and sequences beginning with "upstream" antecedents need elucidation.

CONCLUSION

We call on epidemiologists to continue and even invigorate the debate about future directions. Epidemiology is a discipline dedicated to understanding the causes of health states in populations. Should we do no more than maintain the status quo, our discipline stands to lose its central role as a science of public health. The advent at the macrolevel of information systems and at the microlevel of molecular genetics threatens to split the discipline and even to divorce it from public health. As we have noted elsewhere, the risk factor paradigm does not command the range and depth to avert such a split (70).

Other recent writings express some similar concerns. The paradigm urged here will place epidemiology in an integrative role by explicitly incorporating levels of organization and process. We must allow, however, that Kuhn would not have conceded that our advocacy is toward a paradigm shift. Far from discarding past achievements, this paradigm aims to make use of the contributions of each of the preceding eras of epidemiology, including the risk factor paradigm. Here we add that the new paradigm also preserves the contribution of the miasmatists by restoring the historic interest of the discipline in the broad environmental causation of disease. From the germ theorists, moreover, the paradigm draws lessons about the application of microlevel discoveries, now at the molecular level, to understand and improve population health.

This paper goes no further than sketching a framework. In terms of methods, we indicate only some areas that need development. Epidemiologic designs suited to the study of dynamic and multilevel processes still wait to be elaborated (70, 77). Here we may learn from other disciplines. If epidemiologists are to preserve their own public health tradition, however, we believe they will have

to make a conscious choice to broaden as well as to deepen their approach to the health problems of the future.

Literature Cited

1. Alameida-Filho N. 1991. The epistemological crisis of contemporary epidemiology: paradigms in perspective. *Sante Culture Health* 8:145–66
2. Barber B. 1961. Resistance by scientists to scientific discovery. *Science* 134:596–602
3. Beaglehole R, Bonita R. 1997. *Public Health at the Crossroads: Achievements and Prospects.* Cambridge, UK: Cambridge Univ. Press. 200 pp.
4. Bennett LA, Ames GM. 1985. *The American Experience with Alcohol: Contrasting Cultural Perspectives.* New York: Plenum. 490 pp.
5. Blaser MJ. 1998. Helicobacters are indigenous to the human stomach: duodenal ulceration is due to changes in gastric microecology in the modern era. *Gut.* 43:721–7
6. Buck C. 1974. Popper's philosophy for epidemiologists. *Int. J. Epidemiol.* 4:159–68
7. Cassel J. 1964. Social science theory as a source of hypotheses in epidemiologic research. *Am. J. Public Health* 54:1482–88
8. Cassel J. 1974. Psychosocial processes and "stress": theoretical formulation. *Int. J. Health Serv.* 4:471–81
9. Cassell EJ, Lebowitz MD. 1976. Causality in the environment and health: the utility of the multiplex variables. *Perspect. Biol. Med.* 19:338–43
10. Chiu AO, Bayliss D. 1995. Black box epidemiology. *Epidemiology* 6:464
11. Clayton D, Hills M. 1993. *Statistical Models in Epidemiology.* New York: Oxford Univ. Press. 376 pp.
12. Cook TD, Shadish WR. 1994. Social experiments: some developments over the past fifteen years. *Annu. Rev. Psychol.* 45:545–80

12a. Doll R. 1998. Uncovering the effects of smoking: historical perspectives. *Stat. Meth. Med. Res.* 7:87–117

12b. Diez-Roux A. 1998. Bringing context back into epidemiology: variable and fallacies in multilevel analysis. *Am. J. Public Health* 88:216–22

13. Evans A. *Causation and Disease: a Chronological Journey.* New York: Plenum. 238 pp.
14. Eyler JM. 1979. *Victorian Social Medicine: the Ideas and Methods of William Farr.* Baltimore, MD: Johns Hopkins Univ. Press. 262 pp.
15. Feinstein AR. 1988. Scientific standards in epidemiologic studies of the menace of daily life. *Science* 242:1257–63
16. Fleck L. 1935. *Genesis and Development of a Scientific Fact.* Chicago: Univ. Chicago Press. 203 pp.
17. Gordis L. 1988. Challenges to epidemiology in the next decade. *Am. J. Epidemiol.* 128:1–9
18. Greenland S. 1987. Interpretation and choice of effect measures in epidemiologic analysis. *Am. J. Epidemiol.* 125:761–68

18a. Greenland S. 1998. Induction versus Popper: substance versus semantics. *Int. J. Epidemiol.* 27:543–48

19. Haila Y, Levins R. 1992. *Humanity and Nature: Ecology, Science and Society.* London: Pluto. 270 pp.

19a. Halloran ME, Struchiner CJ. 1991. Study designs for dependent happenings. *Epidemiology* 2:331–38

20. Haynes SN. 1992. *Models of Causality in Psychopathology: Toward Dynamic, Synthetic and Nonlinear Models of Behavior Disorders.* New York: Macmillan. 260 pp.
21. Hill AB. 1965. The environment and disease: association or causation? *Proc. R. Soc. Med.* 58:295–300
22. Deleted in proof
23. Kelsey JL, Whittemore AS, Evans AS, Thompson WD. 1996. *Methods in Observational Epidemiology.* New York: Oxford Univ. Press. 432 pp.
24. Kendler KS, Karkowski LM, Walsh D. 1998. The structure of psychosis: latent class analysis of probands from the Roscommon Family Study. *Arch. Gen. Psych.* 55:502–4
25. Khoury MJ, Yang Q. 1998. The future of genetic studies of complex human

disease: an epidemiologic perspective. *Epidemiology* 9:350–54
26. Koopman JS, Longini IM, Jacquez JA, Simon CP, Ostrow DG, et al. 1991. Assessing risk factors for transmission of infection. *Am. J. Epidemiol.* 133:1199–1209
27. Kramer MS, Lane DA. 1992. Causal propositions in clinical research and practice. *J. Clin. Epidemiol.* 45:639–49
28. Krieger N. 1994. Epidemiology and the web of causation: has anyone seen the spider? *Soc. Sci. Med.* 39:887–903
29. Kuh D, Ben-Shlomo Y. 1997. *A Life Course Approach to Chronic Disease Epidemiology.* New York: Oxford Univ. Press. 336 pp.
30. Kuhn TS. 1962. *The Structure of Scientific Revolutions.* Chicago: Univ. Chicago Press. 226 pp.
31. Lambert R. 1963. *Sir John Simon and English Social Administration.* London: Macgibbon & Kee. 669 pp.
32. Levins R. 1997. When science fails us. *Forests, Trees and People Newsl.* 32/33:15–25
33. Levins R, Lewontin R. 1987. *The Dialectical Biologist.* Cambridge, MA: Harvard Univ. Press. 336 pp.
34. Lilienfeld AM, Lilienfeld DE. 1982. Epidemiology and the public health movement: a historical perspective. *J. Public Health Pol.* 3:140–49
35. Loomis D, Wing S. 1990. Is molecular epidemiology a germ theory for the end of the twentieth century? *Int. J. Epidemiol.* 19:1–3
36. Mackenbach JP. 1995. Public health epidemiology. *J. Epidemiol. Community Health* 49:333–34
36a. Macmahon B, Pugh T, Ipsen J. 1960. *Epidemiologic Methods.* Boston: Little Brown. 361 pp.
36b. Markel H. 1997. *Quarantine! East European Jewish Immigrants and the New York City Epidemics of 1892.* Baltimore, MD: Johns Hopkins Univ. Press. 261 pp.
37. Marshall BJ, Armstrong JA, McGechie DB, Glancy RJ. 1983. Attempts to fulfil Koch's postulates for pyloric campylobacter. *Med. J. Australia* 142:436–43
38. Marshall BJ, Warren JR. 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1:1311–15
39. Maxcy KF. 1941. *Papers of Wade Hampton Frost, MD: a Contribution to Epidemiological Method.* New York: Commonwealth Fund. 628 pp.
40. McMichael A. 1998. London's McMichael previews his keynote SER address. *Epimonitor* 19:1–7
41. Meyer-Bahlburg H, Ehrhardt AA. 1980. Neurobehavioral effects of prenatal origin: sex hormones. *Prog. Clin. Biol. Res.* 36:93–107
42. Miettinen OS. 1985. *Theoretical Epidemiology.* New York: Wiley. 359 pp.
43. Nijhuis HGJ, Van der Maesen LJG. 1994. The philosophical foundations of public health: an invitation to debate. *J. Epidemiol. Community Health* 48:1–3
44. Ottman R. 1994. Epidemiologic analysis of gene-environment interaction in twins. *Genet. Epidemiol.* 11:75–86
45. Pearce N. 1996. Traditional epidemiology, modern epidemiology, and public health. *Am. J. Public Health* 86:678–83
46. Perrera FP. 1997. Environment and cancer: who are susceptible? *Science* 278: 1068–73
47. Riley MW. 1963. *Sociological Research: a Case Approach.* New York: Harcourt, Brace & Work. 753 pp.
48. Risch N, Merikangas K. 1996. The future of genetic studies of complex human disease. *Science* 273:1516–17
49. Rose G. 1992. *The Strategy of Preventive Medicine.* New York: Oxford Univ. Press. 160 pp.
50. Rosen G. 1975. *Preventive Medicine in the United States 1900–1975: Trends and Interpretations.* New York: Prodist. 94 pp.
51. Ross R. 1911. *The Prevention of Malaria.* London: Murray. 2nd ed.
52. Rothman KJ. 1976. Causes. *Am. J. Epidemiol.* 104:587–92
53. Rothman KJ. 1981. The rise and fall of epidemiology, 1950–2000 A.D. *N. Engl. J. Med.* 34:600–2
54. Rothman KJ, ed. 1988. *Causal Inference.* Boston: Epidemiol. Resources. 207 pp.
55. Rothman KJ, Greenland S. 1998. *Modern Epidemiology.* Philadelphia: Lippincott-Raven. 737 pp.
56. Ryle JA. 1936. *The Natural History of Disease.* London: Oxford Univ. Press. 484 pp.
57. Ryle JA. 1948. *Changing Disciplines: Lectures on the History, Method, and Motives of Social Pathology.* London: Oxford Univ. Press. 122 pp.
58. Savitz DA. 1994. In defense of black box epidemiology. *Epidemiology* 5:550–52
59. Scamman CL. 1934. *Papers of Charles V. Chapin, M.D.* New York: Commonwealth Fund. 244 pp.
60. Schnall PL, Kern R. 1981. Hypertension in American society: an introduction to historical materialist epidemiology. In

The Sociology of Health and Illness, ed. P Conrad, R Kern, pp. 97–122. New York: St. Martin's

60a. Schwartz S. 1994. The fallacy of the ecologic fallacy: the potential misuse of a concept and the consequences. *Am. J. Public Health.* 84:819–29

61. Shy C. 1997. The failure of academic epidemiology: witness for the prosecution. *Am. J. Epidemiol.* 145:479–84

62. Skrabanek P. 1994. The emptiness of the black box. *Epidemiology* 5:553–55

63. Sommer A. 1997. Clinical research and the human condition: moving from observation to practice. *Nat. Med.* 3:1061–63

64. Stallones R. 1980. To advance epidemiology. *Annu. Rev. Public Health* 1:69–82

65. Susser E, Brown A, Matte T. 1998. Prenatal antecedents of neuropsychiatric disorder over the life courses: collaborative studies of US birth cohorts. In *Childhood Onset of "Adult" Psychopathology: Clinical and Research Advances,* ed. J Rappoport. New York: Am. Psychiatr. Press. In press

65a. Susser M. 1973. *Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology.* New York: Oxford Univ. Press. 181 pp.

66. Susser M. 1982. Period effects, generation effects and age effects in peptic ulcer mortality. *J. Chron. Dis.* 35:29–40

67. Susser M. 1987. Falsification, verification and causal inference in epidemiology: reconsiderations in the light of Sir Karl Popper's philosophy. In *Epidemiology, Health and Society; Selected Papers,* ed. M Susser, pp. 82–93. New York: Oxford Univ. Press

68. Susser M. 1991. What is a cause and how do we know one? A grammar for pragmatic epidemiology. *Am. J. Epidemiol.* 133:635–48

68a. Susser M. 1994. The logic in ecological: 1. The logic of analysis. II. The logic of design. *Am. J. Public Health* 84:825–29; 830–35

68b. Susser M. 1998. Does risk factor epidemiology put epidemiology at risk? Peering into the future. *J. Epidemiol. Community Health* 52:608–11

69. Susser M, Susser E. 1996. Choosing a future for epidemiology: I. Eras and paradigms. *Am. J. Public Health* 86:668–73

70. Susser M, Susser E. 1996. Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. *Am. J. Public Health* 86:674–77

71. Susser M, Stein Z. 1962. Civilization and peptic ulcer. *Lancet* 1:115–19

72. Taubes G. 1995. Epidemiology faces its limits. *Science* 269:164–69

73. Taubes G. 1998. As obesity rates rise, experts struggle to explain why. *Science* 280:1367–68

74. Terris M. 1964. *Goldberger on Pellagra.* Baton Rouge: Louisiana State Univ. Press. 395 pp.

75. Terris M. 1987. Epidemiology and the public health movement. *J. Public Health Pol.* 8:315–29

76. Vena JE, Vana JE, Trevisan M. 1997. The legacy of epidemiology in the Department of Social and Preventive Medicine: a commemoration of the sesquicentennial of the State University of New York at Buffalo School of Medicine and Biomedical Sciences. *Am. J. Epidemiol.* 146:891–95

77. Von Korff M, Koepsell T, Curry S, Diehr P. 1992. Multi-level analysis in epidemiologic research on health behaviors and outcomes. *Am. J. Epidemiol.* 135:1077–82

78. Waldrop MM. 1992. *Complexity: the Emerging Science at the Edge of Order and Chaos.* New York: Simon and Schuster. 384 pp.

79. Weed DL. 1988. Beyond black box epidemiology. *Am. J. Public Health* 88:12–14

79a. Weinberg S. 1998. The revolution that didn't happen. *NY Rev. Books* 45:48–52

80. Wing S. 1994. Limits of epidemiology. *Med. Global Survival* 1:74–86

81. Winslow CA. 1929. *The Life of Hermann M. Biggs, M.D., D.Sc., LL.D.: Physician and Statesman of the Public Health.* Philadelphia: Lea & Febiger. 432 pp.

82. Winslow CA. 1943. *The Conquest of Epidemic Disease: A Chapter in the History of Ideas.* Madison, Wisc.: Univ. Wisconsin Press. 411 pp.

82a. Wolf S, Wolff HG. 1943. *Human Gastric Function.* New York: London: Oxford University Press

83. Wynder EL. 1997. Tobacco as a cause of lung cancer: some reflections. *Am. J. Epidemiol.* 146:687–94

84. Zito G. 1975. *Methodology and Meanings: Varieties of Sociological Inquiry.* New York: Praeger. 175 pp.

85. Zuckerman H. 1988. The sociology of science. In *Handbook of Sociology,* ed. NJ Smelser, pp. 511–74. Newbury Park, CA: Sage