Prisoners of the Proximate: Loosening the Constraints on Epidemiology in an Age of Change

ANTHONY J. MCMICHAEL


Published with permission of OUP ©
American Journal of EPIEMIOLOGY

Volume 149
Number 10
May 15, 1999

COMMENTARY

Prisoners of the Proximate: Loosening the Constraints on Epidemiology in an Age of Change

A. J. McMichael

"Modern epidemiology" has a primary orientation to the study of multiple risk factors for chronic noncommunicable diseases. If epidemiologists are to understand the determinants of population health in terms that extend beyond proximate, individual-level risk factors (and their biological mediators), they must learn to apply a social-ecologic systems perspective. The mind-set and methods of modern epidemiology entail the following four main constraints that limit engagement in issues of wider context: 1) a preoccupation with proximate risk factors; 2) a focus on individual-level versus population-level influences on health; 3) a typically modular (time-windowed) view of how individuals undergo changes in risk status (i.e., a life-stage vs. a life-course model of risk acquisition); and 4) the, as yet, unfamiliar challenge of scenario-based forecasting of health consequences of future, large-scale social and environmental changes. The evolution of the content and methods of epidemiology continues. Epidemiologists are gaining insights into the complex social and environmental systems that are the context for health and disease; thinking about population health in increasingly ecologic terms; developing dynamic, interactive, life-course models of disease risk acquisition; and extending their spatial-temporal frame of reference as they perceive the health risks posed by escalating human pressures on the wider environment. The constraints of "the proximate" upon epidemiology are thus loosening as the end of the century approaches. Am J Epidemiol 1999;149:887–97.

epidemiology; history; methods; population; social environmental

The past 3 decades have witnessed the methodological consolidation of "modern epidemiology," with its particular orientation to studying the multiple risk factors for chronic noncommunicable diseases. That conceptual and methodological orientation arose from midcentury as epidemiologists formally engaged in the study of diseases of long latency, multiple causality, and apparently noninfectious etiology (1). (I will refer to such diseases as "chronic diseases," while also noting the likely involvement of infectious agents in the etiology of some of them.)

It is axiomatic that the theoretic framework within which we formulate our research questions determines the scope, content, and social relevance of our answers. The question of context has excited much recent debate about the mission, models, and methods of modern epidemiology (2). To understand the determinants of population health in terms beyond proximate, individual-level risk factors (and their biological mediators) requires a social-ecologic systems perspective. Yet, modern epidemiology has largely ignored these issues of wider context. We have typically assumed that populations are merely aggregates of free-range individuals and that methodologically correct local studies can estimate presumed universal individual-level risk relations. Such an approach, however, forfeits understanding of the causes and distribution of disease within populations and thus restricts the social usefulness of the research, particularly in a rapidly changing world.

That familiar word, "understanding," has great epistemologic significance. It refers to the intellectual framework within which we gather and interpret our
observations. Figure 1 shows how a Dutch artist of the sixteenth century, Cornelis de Jode, visualized an Australian kangaroo by a sixteenth-century Dutch artist (Cornelis de Jode), based on descriptions by early explorers.

PHASES IN THE HISTORY OF EPIDEMIOLOGY

Was there a time when epidemiologists took a more deliberate interest in the social patterning of disease? Table 1 summarizes the historical ebb and flow of epidemiologic ideas about the determinants of population health and disease over recent centuries.

The first stirrings of formal epidemiology were in John Graunt's descriptive analysis of London's bills of mortality in the 1660s (8). At that time, the classical-rationalist approach to the causes of illness was being challenged by the ontologic notion of specific diseases (9). The idea arose that empirical observations could usefully be made about disease causation (such as in Ramazzini's linking of particular occupations with particular diseases (10)).

In the late 1700s, epidemiologic enquiry in Britain began to address several specific problems. Thomas Lind carried out his miniaturized controlled trials of the dietary prevention of seaboard scurvy. Case series of unusual diseases—such as scrotal cancer in chimney sweeps and colic in Devon cider drinkers—indicated particular environmental causes. However, analyses of differential mortality and disease causation were still principally framed in terms of geography, relative wealth, and occupation. Epidemiologic concepts remained broad; methods of inquiry were, by today's standards, extremely crude.

In early nineteenth-century industrializing Europe, a more humane, egalitarian, post-Enlightenment interest emerged in public health problems associated with urban squalor, deprivation, and crowding. The observations of Villermé in France and Virchow in Germany emphasized the association of diseases with urban conditions, poverty, and hazardous occupations (11, 12). In 1845, Engels described Manchester's factory-working masses as "pale, lank, narrow chested, hollow-eyed ghosts" (13, p. 128) afflicted with rickets and scrofula. He documented a doubling in death rates between families living in the best and the worst housing.

Throughout the first half of the nineteenth century, the miasma theory of disease causation prevailed, tempered by minority support for the alternative idea of "contagion," first mooted in the sixteenth century (14).
The manifest tendency of certain diseases such as influenza, cholera, and yellow fever to sweep over entire populations seemed to confirm that nonspecific miasmatic emanations from decaying organic matter were the cause. Accordingly, William Farr explained the marked district variations in cholera mortality within London in terms of housing quality, residential height above sea level, and general air quality (15). This was epidemiologic research that moved easily between population, neighborhood, and individual levels (1). Overall, a social and environmental holism in public health characterized the early 1800s.

Then, from midcentury in Britain, the application of Chadwick’s “Sanitary Idea” foreshadowed a narrower, technical stratagem to banish miasmas by purifying the water and air. Hamlin (16) described this rise of sanitary engineering as the “degreening” of nineteenth century public health. Meanwhile, the accruing evidence of contagion, especially from Semmelweis’ studies of puerperal fever (17) and Snow’s studies of cholera (18), challenged the miasma theory. Eventually, in the 1880s, the germ theory reoriented epidemiology toward the idea of specific causation: Diseases could each be understood in terms of a single causal infectious agent.

This powerful idea was reinforced by the discovery of certain occupational exposures as causes of cancer and by the implication of specific vitamin deficiencies in nutritional disorders. Meanwhile, early crude ideas about the human genetic determinants of individual constitutional “fitness” were boosted by the rediscovery of Mendel’s work on particulate genetics in peas (19). So, here, in the early twentieth century was a world in which disease was caused by germs, carcinogens, vitamin deficiencies, and genes. Disease causation could now be interpreted in terms of proximate personal exposures and attributes.

However, epidemiologists (like humans everywhere) keep returning to the search for wider meaning and understanding. In the second quarter of this century, the triumphant germ theory was broadened to accommodate the interactive roles of environmental conditions and host susceptibility. In the United States in 1927, Wade Hampton Frost sought to “establish a theory” to explain the distribution of disease within populations (20). Greenwood (21) in England and Sydenstricker (22) in the United States wrote of “crowd diseases,” the social environment, and social inequalities in health.

Since the 1950s, as infectious diseases receded, epidemiologists in developed countries have been preoccupied with chronic diseases of complex etiology. Faced with this diversity of diseases and risk factors, they adopted an essentially empirical approach. Numeric reasoning, based on statistical modeling, has been central to the effort. A growing preoccupation with the role of multiple proximate risk factors largely eclipsed ideas of social causation of disease. We have thus evolved a modern epidemiology that is adept at
determining which individuals are at increased risk, but not at understanding disease distribution within and between populations. We have been busy reacting to our consumer society’s procession of new, potentially hazardous exposures: mobile telephones, replacement estrogens, vitamin supplements, mad cows, photochemical smog, and endless new chemicals and drugs.

The explanation goes deeper, however. As also followed the rise of the germ theory, modern epidemiology’s search for specific proximate causes has deflected us from social-contextual models of disease causation. It has changed the conceptual framework. Hence, the recent calls for restoring the population perspective, for recognizing that population history, culture, and social structure determine the level and internal distribution of disease risk (4, 5, 23, 24). Some argue that a theoretical “black box” positivist epidemiology should be recast within an “ecologic” paradigm attuned to population processes, interdependencies, and multi-level causality (1, 3). This proposed systems perspective is being reinforced by emerging ideas about the life-course evolution of chronic disease risk (25).

So epidemiology is edging, again, toward a more integrative view of the sources and distribution of disease. This time, we are countenancing a more ecologic view of how the social and environmental conditions of life influence population health. (Epidemiologists have long misused this word “ecologic,” borrowed from sociology (26). Ecology is the formal study of the interrelations between groups of organisms, populations, and species and their surroundings. I here use the word liberally to refer to contexts characterized by interdependencies between individuals and groups of humans and their environments.) In seeking an epidemiology that integrates across macro-, meso-, and microlevels of causal analysis and deals with complex social and ecologic relations, we must understand the constraints that currently confine us.

“PROXIMATE” CONSTRAINTS ON MODERN EPIDEMIOLOGY

The mind-set and methods of modern epidemiology entail four types of proximate constraints upon our research agenda. They are:

1. Our preoccupation with proximate risk factors
2. Our focus on individual-level versus population-level influences on health
3. Our typically modular (time-windowed) view of how individuals undergo changes in risk status
4. The, as yet, unfamiliar challenge of scenario-based forecasting of health consequences of future, large-scale, social and environmental changes.

Figure 2 represents these four interrelated constraints. Epidemiology, shown as a little orb, is “imprisoned” at bottom left. Let us consider these four constraints in more detail.

Preoccupation with proximate risk factors

Most contemporary epidemiologic research focuses on proximate risk factors, that is, specific exposures, circumstances, or behaviors that are measurable at the individual level. During 5 decades, this has yielded many important new insights into chronic disease causation. With seemingly fewer large effects remaining to be discovered, however, such research is becoming less informative (1, 27, 28). Nevertheless, important findings continue to be made (29), in part reflecting improved measurement methods and the use of very large studies or meta-analyses (8, 30).

The preoccupation with seeking specific proximate risk factors has been disdainfully referred to as “risk factorology” (28, 31). The criticism is misleading; the fault is not in doing such studies, but in only doing

![Figure 2](image-url)
such studies. We should also be looking upstream for a fuller account of disease causation within a population context; we must extend our focal length. Consider alcohol consumption and liver cirrhosis. The proximate “cause” of liver cirrhosis is the individual’s drinking behavior—or the ethanol itself. However, we also wish to account for the within-population distribution of alcohol-related liver cirrhosis. Are there differences between generations, social classes, ethnic groups, or occupations? Are those differences due to subculture, peer pressure, commercial advertising, opportunity, or employment status?

How epidemiologists relate to that upstream category of question is influenced by whether we perceive the distal determination of disease as being part of either a linear causal chain or a systems-based causal web. The former model may tempt us to argue that, while it is the epidemiologist’s task to define the downstream relation between alcohol consumption and disease, it is the separate task of social scientists working upstream to elucidate the determinants of drinking behavior and why it differs between population subgroups (32). The latter, systems-based, approach envisions a causal web that extends inward, via multiple paths, from the encircling realms of the population’s history, culture, and socioeconomic relations, through residential conditions and subpopulation attitudes, to the inner proximate factors of individual behaviors and exposures and their biomedical manifestations (33). Causal processes within this web are not necessarily linear and sequential, but may involve interactions and feedbacks. We may not, as epidemiologists, like this complexity, but life is like that; it comprises complex systems.

Attuning epidemiologic methods to social complexity. Kaufman and Cooper argue that “the traditional epidemiologic method is ill-suited for considering social quantities as risk factors” (34). We should be seeking a systems-based model of the observable world, they say, and not statistically estimating risks by reference to what, in this context, entails an untenable (unobservable) “counterfactual.” For example, we can consider a smoker who, we can imagine, does not smoke (i.e., the idealized unexposed “counterfactual”) as being otherwise still the same person (therby guiding our selection of an appropriate unexposed comparison group). However, we cannot consider that a socially disadvantaged person, if no longer socially disadvantaged, would otherwise be the same person, since many other complex determinants and concomitants of socioeconomic status would also be changed.

With race and gender, they argue, the counterfactuals become even more meaningless. If we seek to explain the differences in health status between Black and White Americans in proximate biomedical terms, then we might choose to treat socioeconomic status as a potential confounder that requires routine adjustment, like age and sex. However, this naively assumes that the “risk” attributes of socioeconomic status are independent of those associated with race, including the experience of racism at individual, institutional, and societal levels (35). Kaufman et al. (35) therefore advocate more creative approaches to causal explanation in social epidemiology, drawing on the concepts of infectious disease epidemiology and systems analysis. Using multilevel or pathway modeling, and collaborating with other disciplines, epidemiologists can then develop quantitative and structural analyses of how social variables affect health outcomes.

Consider the widening social class gradient in coronary heart disease (CHD) mortality in Britain over the past 2 decades (36, 37). This divergence reflects a much sharper decline in CHD deaths in the upper classes than in the lower classes. The principal challenge is to understand why, in terms of material circumstances, social relations, knowledge, and behaviors, this class-related gap in the risk of dying from CHD has widened (38). Why is the class-related CHD gradient evident in northern European countries but not in France, Switzerland, and Mediterranean–Europe countries, where, instead, the class gradient is most evident for cancers and gastrointestinal diseases (39)? Explanations for such patterns cannot come from exclusively downstream research. From the experience of social epidemiology, we need to learn how to conceptualize and analyze this complex of proximal and distal influences on health outcomes (1, 33).

Epidemiologists must, of course, continue to study proximate causes. Further important causal relations will be revealed at that level. However, if that is all we do, ignoring wider social-environmental causes, then we risk reinforcing inefficient programs of local “health promotion” accessible only to the better-educated, better-resourced stratum of society (40). Effective and equitable social interventions require an understanding of the contextual determinants of health risk distribution within the population.

Rose’s “sick population” perspective. There is another, related issue. We may wish to distinguish between factors that explain the occurrence of disease in individuals within a specific population and those that account for the overall population rate of the disease. Geoffrey Rose (41) pointed out that a factor that is an important cause of disease in individuals within a population may differ from one that primarily determines the disease rate within the whole population.

Rose illustrated his argument with risk factors from the familiar downstream catalogue, each of which
could be meaningfully measured at either the individual or the population level. Consider an example in which population A has a higher rate of CHD than does population B. Within each population, individuals who smoke have twice the CHD risk of nonsmokers. Nevertheless, the prevalence of smoking is identical in each population and, hence, cannot account for the between-population difference in CHD rates. However, population A has a generalized higher consumption of saturated fat than does population B. Hence, the between-population difference in CHD rates is explained by differences in diet, not smoking, even though smoking is the dominant risk factor at the individual level.

That example would be equally plausible if we interchanged smoking and dietary fat. Often, however, the factors that operate at individual and group levels are not interchangeable, being of qualitatively different scale and content. This brings us to the second category of constraint.

Levels of causal analysis: individual versus population

Implicit in the focus on proximate risk factors is an assumption that the individual is the site of etiologic action. We thus assume that, while social and environmental influences may originate on a broader front upstream, they ultimately become manifest, as risk factors, at the level of the individual. The individual, therefore, is viewed as the natural unit of epidemiologic observation. Larger-scale variables that affect whole groups or populations, such as poverty and cultural disruption, are only important because they translate into individual-level risk factors. Poverty affects diet, cultural disruption breeds alcoholism, and so on.

This population/individual distinction, however, needs careful consideration. Are we, as in the previous section, merely distinguishing between upstream social contexts and their downstream proximate manifestations? Or is there a category of risk factor that, in some collective way, influences the health of the population at large via processes that have no direct downstream manifestation? Further, complex entities such as poverty, for example, can have very different meanings and can measure qualitatively different constructs at the individual and population levels (42).

Susser has recently reminded us that the subject matter of epidemiologic research is "ecological in the original biological sense of organisms in a multilevel interactive environment" (1, p. 609). The sociologist Émile Durkheim, in the late nineteenth century, was one of the first to argue that society's characteristics at large can affect a health outcome (43). The suicide rate, he said, was not simply the aggregate of individual suicidal tendencies. Rather, it reflected properties of the population: the underlying social values, the pattern of social relations, and the moral significance accorded to the act of suicide. The usual individualistic accounts, he said, were therefore deficient.

Consider the well-known example of herd immunity in relation to infectious disease (44). If, within a population, sufficient of the potential contacts of a primary infectious case are immune, then the average number of secondary cases will be too few for the infection to spread. By definition, it is only a population that can have herd immunity; it has no direct, corresponding representation at the individual level. It is this constitutional property of the population that determines whether the infectious disease can be sustained within the population. Therefore, there is an interdependence of risk between the unit members of the population, reflecting the prior experience of that population as a whole.

Indeed, infectious disease epidemiologists well understand that the individual-level risk of infectious disease often reflects population-level characteristics (45). For example, the sexual spread of human immunodeficiency virus within a population is strongly determined by the pattern of sexual activity within the population, which, in turn, reflects economic relations, demographic mobility, and cultural traditions. An individual's risk, for a given number of sex partners, depends on who is having sex with whom and on the prevalence of the infection within that contact network. Individuals have sex; populations have patterns of sex. Both are risk factors, but at different levels.

The distribution of income within a population is another interesting population characteristic: Individuals have a personal income; populations have an income distribution. At the individual level, income has a well-known relation to health: Poor individuals die younger. Davey-Smith et al. (46, 47) have recently demonstrated a clear gradient in mortality across 14 income classes in US Blacks and Whites separately. Low absolute personal income, we presume, affects individual health principally via material deprivation.

At the population level, however, the relation between wealth and health is different. Wilkinson (48) and others have demonstrated, in developed countries, that differences in average life expectancy between countries show little correlation with average income. Rather, average life expectancy correlates inversely with the extent of within-population income inequality. This relation reflects health experiences across the income scale and is not a simple arithmetic consequence of extremely poor health in a deprived minority. Nor is it explained by interpopulation differences in levels of public expenditure on health and welfare (49).
At a subnational scale, Kaplan et al. (49) have shown for the US states that the inverse correlation between average state income and mortality is entirely accounted for by differences in within-state income inequality.

Something about a population's internal income relations thus appears to affect overall population health independently of any effect of average personal income. Recently, in a study in 39 US states, Kawachi et al. (50) showed that, in states where income differences are greater, people experience their social environment as more hostile; they are less likely to join community organizations and more likely to mistrust other people. Those researchers conclude that the gap between rich and poor affects social organization and that the resulting loss of "social capital" may impair the population's health.

How is this effect mediated? Some of these investigators invoke individual-level physiologically based explanations via the neuroendocrine response to stress deprivation, insecurity, and chronic stress. Studies in primate colonies indicate that individual-level measures of status-related stress, cortisol levels, and atherogenesis are positively associated (51, 52). Recent epidemiologic evidence suggests that the rate of progression of atherosclerosis is inversely proportional to socioeconomic status (53) and occupational reward:effort ratio (54).

So, is this search for individual-level explanation merely biological realism? Or does it reflect the hegemonic power of individual-based biomedical thinking? Might it be that lessened income inequality reduces rates of alienation, violent crime, and infectious disease risk behaviors in ways that confer protection on everyone? Perhaps the generalized relative paucity of material assets (including health care facilities) in unequal societies (55) and, indeed, the legacies of social, cultural, and political history of those unequal societies impair their health prospects (56). We are not well placed to answer these questions because we have not yet undertaken the types of multilevel analysis that could elucidate coexistent and interactive influences at the population and individual levels (33). Perhaps part of the widely reported individual-level inverse association between income and mortality actually results from the group-level experiences of subpopulations experiencing relative deprivation. The debate continues.

The bleak mortality experience of central and eastern Europe in recent decades raises other population-level questions. During the 1970s and 1980s, as life expectancies increased in western Europe, central and eastern European countries experienced increasing mortality from many causes, including heart disease, stroke, respiratory infections, accidents, and violence, all especially in men (57). The striking east-west divergence in death rates during those decades cannot be explained simply in terms of individual-level risk behaviors (such as smoking and drinking) or exposure to environmental pollutants (58). Watson (58) postulates that the generalized state control of daily life, poverty of community networks, and suppression of individual initiative induced a collective learned helplessness and community disengagement, while reinforcing inward-looking traditional family values (59). This loss of social capital eroded the traditional role of men more than that of women and created conditions that jeopardized health. In post-Soviet Russia, social frustration and disintegration appear to have underlain the dramatic surge in premature male mortality attributable to excessive alcohol consumption during 1990-1995 (60). Regional analysis shows that the local decline in life expectancy was clearly correlated with local labor turnover, recorded crime rates, and unequally distributed income (61). Rapid social and economic change apparently disrupts stabilizing social-behavioral patterns, leading to self-destructive behaviors.

There is other historical evidence that social, cultural, and political characteristics of a society influence susceptibility to disease. After European settlement, the emergence of CHD in indigenous Amerindian, Maori, and Australian Aboriginal populations reflected the social structures of both the indigenes and the settler populations (62). More generally, changes in the life expectancy of the native population was influenced by the social and political structures of settler populations and by the form of interaction between populations.

These examples underscore the importance of studying the role of population-level influences on health. Such factors, intrinsically difficult to characterize, are neither the mere aggregation of individual risk factors nor the directly connected upstream determinants of proximate factors. Rather, there appear to be constitutional properties of populations, such as herd immunity, income inequality, and social morale, that affect health processes at a supra-individual level, in addition to any manifestations of risk that they might induce at the individual level.

Life-stage versus life-course models of risk acquisition

A third constraint on our thinking is that much epidemiologic research implies a static, modular view of the acquisition of risk. In caricature, we take 1,000 healthy adults, add 10 cigarettes per day or a regular dose of air pollution, and then see by how much the health status changes. This approach is not necessarily wrong, but it often gives incomplete or misleading answers. Much risk of disease, especially noninfectious disease, evolves over a lifetime via cumulative and interactive processes.

Am J Epidemiol Vol. 149, No. 10, 1999
An intriguing example comes from research done in a rural population in The Gambia, West Africa. Survey data show that the average weight of adult women in that population fluctuates seasonally, being 5–7 percent greater in the harvest season. An historical survival analysis of children born around midcentury revealed that those children born during the harvest season experienced distinctly better survival in adulthood than did those born in the nonharvest season. By the fifth decade of life and with no survival difference apparent before age 15 years, survival in the two groups was approximately 65 and 45 percent, respectively (63). Something to do with perinatal nutrition has profoundly affected long-term biological robustness.

Asthma offers another example. For several decades, we have sought immediate environmental causes: air pollutants, environmental tobacco smoke, and allergens. Yes, many of these exposures probably trigger attacks in susceptible persons. However, why has the disease increased sevenfold in western populations over the last 25 years? Why, in ecologic terms, are we producing successive generations of children with increasing susceptibility to asthma? Perhaps early childhood experiences (infections, contact with environmental bacteria, vaccinations, etc.) affect immune system programming along allergic or nonallergic pathways (64).

Over the past decade, considerable evidence has accrued that biological processes and experiences in early life, especially fetal life, affect lifelong susceptibilities to various adult disease processes: cardiovascular disease, diabetes, immune disorders, respiratory diseases, and others. Research along these lines in Britain, especially by Barker (65) and others, has greatly extended Forsdahl’s work in Norway in the 1970s and, indeed, that of certain British nutritional scientists in the 1930s (66, 67). These ideas need to be subjected to critical tests of replication (68) and require biological corroboration from laboratory and clinical sciences. Further, the public health (“attributable risk”) significance of these potentially important early-life influences on adult disease risks should also be assessed.

Inevitably, many of these dynamic causal models are complex. The relative importance of early-life and adult-life influences varies between categories of disease outcome. A cohort study of Scottish men found that the risk of stroke and stomach cancer mortality depended primarily on childhood socioeconomic conditions, whereas heart disease mortality was influenced by conditions in both childhood and adulthood (69). Early-life experiences may critically affect some heart disease risk factors, such as triglyceride levels, while cholesterol level and blood pressure are affected by both early and adult lives (70). Further, the combination of being a small baby and a large adult—a combination that maximizes the metabolic mismatch between life stages—markedly elevates the risk of hypertension (figure 3) and of cardiovascular disease mortality (71, 72).

A life-course model of disease etiology is thus evolving, distinct from the static “adult lifestyle” model. It posits coexistent, often interactive, chains of biological and social influences that underpin the development of adult disease risk (73). These life course-based insights, intrinsically scientifically interesting, also have important implications for disease prevention, of course. This line of research will require various new data-analytic techniques, such as repeated-measures techniques, multistate modeling, and adaptive genetic algorithms.

**Forecasting health risks under conditions of global change**

The fourth constraint is our confinement, as essentially empirical scientists, to working in the present and recent past tenses. We are not yet well attuned to forecasting future health risks in relation to potentially important future “exposures.” Indeed, the possibility of major adverse systemic environmental change has not previously pressed upon us.

Because of the combined weight of human numbers, economic activity, and technology, we are starting to change the conditions of life on Earth. Such a change will affect the global patterns of human health and disease (1, 74). For the past 2 centuries, epidemiologists have lived and worked in a world in which large-scale, natural life-support systems have not been perceptibly perturbed and weakened. We no longer live in such a world. We must therefore now think beyond the traditional striving for incremental health gains within populations; we must also address the issue of the sustainability of population health against the prospect of a deteriorating natural environment.

The two best-known global environmental changes are the accumulation of heat-trapping greenhouse gases in the lower atmosphere (troposphere) and the depletion of stratospheric ozone caused by ozone-destroying gases. Authoritative international scientific reviews have semiquantitatively characterized the anticipated, mostly adverse health consequences of these global change processes (75, 76). These, however, are only part of a longer list of newly recognized, global and worldwide environmental changes (77). Other major changes include losses of biodiversity, depletion of supplies of freshwater, degradation of food-producing systems on land and at sea, and the worldwide dissemination of persistent organic chemicals, each entailing potentially great risks to human health (78).
Is the role of epidemiology essentially reactive? Is it limited to helping society understand and tidy up its public health messes after they occur and, thus, to reducing the likelihood of recurrence? Hopefully not. Rather, as the scale of humankind's impact on large biophysical systems increases, triggering unfamiliar global-scale environmental changes, epidemiologists should acquire new skills in anticipatory, scenario-based, health risk assessment (79). The role of such assessment is primarily to assist human society to foresee and understand the range of likely consequences of current and emerging economic, social, and environmental trends (80). Only by entering the rapidly developing arena of "futures studies" (81), rich in interdisciplinary challenge, can epidemiologists engage in health risk assessment that can guide the development of proactive policies to constrain these large-scale environmental changes.

Research into how future changes in world climate, in ambient ultraviolet radiation exposure, and in other large-scale environmental changes are likely to influence health risks is still in an early developmental phase. There are two major categories of research needs:

1. Empirical studies into the relation between relevant environmental variations (e.g., meteorologic variables and ultraviolet radiation levels) and human health outcomes. Such studies can serve two purposes: the extension of knowledge about these causal relations and the detection of early health impacts of these environmental changes.

As time passes, the relevance of the latter use increases.

2. Integrated mathematical modeling of the future health outcomes in relation to the forecast scenarios of environmental change (i.e., scenario-based health risk assessment).

A balance is needed between empirical and predictive research. The latter, relying substantially on integrated mathematical modeling, is important for assessing the range of plausible outcomes (for example, geographic shifts in the potential transmissibility of vector-borne infections, changes in regional food security and levels of malnutrition, and increases in skin cancer incidence rates) and also for revealing gaps in knowledge about relations and processes (79). The former, empirical research will fill those gaps and enhance our capacity to forecast future scenario-based health impacts. Empirical studies of the health consequences of recent variations in exposure—for example, El Niño climatic episodes (82)—will yield further understanding of these environment-health relations.

The advent of these macroscale environmental risks to human health means that a future-oriented interdisciplinary research effort is required in which epidemiologists play a substantive role.

CONCLUSIONS

As we enter a new century, we epidemiologists must broaden our causal models and recognize the important ecologic dimensions of social-environmental influences.
on health and disease. Last century, epidemiologists recognized that the unhygienic conditions and socioeconomic disparities of urban-industrial life were the major cause of disease. Our modern preoccupation has been to understand disease occurrence in individuals in terms of consumer behaviors, individual exposures, metabolic factors, and genes. The advent, early next century, of “bar-coded” individual genotypes on microchips may yet further distract us from the task of managing our social and natural environments.

The landscape is changing, however. Infectious diseases continue their apparent reemergence (83). There are increasing pressures on the wider environment from burgeoning human numbers and economic activity. We are gaining new insights into the complex social and environmental systems that are the context for human health, thinking more about population health in ecologic terms, and extending our spatial-temporal frame of reference. The constraints of “the proximate” upon epidemiology are thus being loosened as we approach century’s end.

ACKNOWLEDGMENTS


Many of the ideas in this paper have evolved from discussions with David Leon, with whom the author teaches a Masters-level unit entitled “Emerging Themes in Epidemiology.” We have each benefited from discussions with our student classes. Helpful comments were also received from George Davey-Smith, Alistair Woodward, and John Lynch.

REFERENCES

17. Semmelweis IP. The etiology, the concept and prophylaxis of childbed fever. Published in 1861. Translated and republished in Medical Classics 1941;5:350–773.
37. Phillimore P. Beattie A, Townsend O. Widening inequality of

Am J Epidemiol Vol. 149, No. 10, 1999

Published with permission of OUP ©
Prisoners of the Proximate: Loosening the Constraints on Epidemiology in an Age of Change

ANTHONY J. MCMICHAEL