REFOCUSING THE LENS: EPIDEMIOLOGIC TRANSITION THEORY, MORTALITY DIFFERENTIALS, AND THE AIDS PANDEMIC

DANIEL S. GAYLIN and JENNIFER KATES
The Lewin Group, 9302 Lee Highway, Suite 500, Fairfax, VA 22031, U.S.A

Abstract—The epidemiologic transition theory presented first by Omran (Omran, A. R. (1971) The epidemiologic transition: a theory of the epidemiology of population change, Milbank Quarterly 49(4), 509–538) was designed to explain global trends in the dynamic relationship between epidemiological phenomena and demographic change. This paper argues that universalizing this theory only partially serves to explain mortality declines over the last century and eclipses key epidemiologic differences between population subgroups based on socioeconomic status, race, and sex. This paper examines morbidity and mortality differentials between population subgroups and demonstrates important inconsistencies with the optimistic trends implied by the epidemiologic transition theory, an argument further developed using the HIV/AIDS pandemic as a case study. The paper argues that these differences should be brought from margins to center to present a more complex and comprehensive picture of how population subgroups experience epidemiologic transitions differently. Copyright © 1997 Elsevier Science Ltd

Key words—epidemiologic transition, mortality differentials, AIDS, infections

INTRODUCTION

...[I]f we are to understand our contemporary reaction to a traditional stimulus, we must distinguish between the unique and the seemingly universal, between this epidemic at this time and this place and the way in which communities have responded to episodic outbreaks of culminating infectious diseases in the past. We have become accustomed in the last half century to thinking of ourselves as no longer subject to the incursions of such ills; death from acute infectious diseases has seemed—like famine—limited to the developing world. Life-threatening infectious ills had become almost by definition, amenable to therapeutic or prophylactic intervention (Rosenberg, 1989).

Extensive worldwide spread of HIV started in the mid to late 1970s. In less than two decades—during the first of which it was unknown and unsuspected—HIV became the first modern pandemic (Mann et al., 1992).

During the past century, the world has experienced dramatic mortality declines. Demographers, anthropologists, and sociologists have attributed global mortality declines to the near disappearance of infectious diseases and their related epidemics. In 1971, Omran published the “The epidemiologic transition: a theory of the epidemiology of population change” in an attempt to capture the changing nature of epidemics and their demographic effects. The major contribution of this theory is its connection between modernization, socioeconomic progress, and mortality decline via changing epidemiologic conditions, primarily the decline in infectious disease. In addition, the epidemiologic transition theory examines mortality differentials between subgroups of the population—particularly between U.S. blacks and whites and between men and women—and argues that these differentials converge over time (actually, the theory favors women slightly over men in terms of mortality gains). Thus, the epidemiologic transition theory has been generalized as a broad explanatory approach, one which is utilized to describe a general global epidemiologic experience.

In this paper, we argue that universalizing the epidemiologic transition theory eclipses key epidemiologic differences between population subgroups, and that these differences should be brought from the margins to center to create a more complete and accurate representation of population morbidity and mortality. When this is done, the picture presented by the epidemiologic transition theory significantly expands to reflect numerous epidemiologic transitions intricately connected to socioeconomic, sex, and racial differences. Furthermore, we argue that mortality trends since the publishing of the last update of the epidemiologic transition theory (Omran, 1983) reflect a troubling rise in infectious diseases—particularly tuberculosis and AIDS—and an increasing disproportionate impact of these epidemics on particular population subgroups. We believe that such an examination and rethinking of the theory not only provides a needed critique, but it enhances the theory’s ability to explain the dynamic relationship between population change and epidemiology as this relationship is experienced by different subgroups of the population. Thus, rather than reject the theory outright,
as some might argue, we believe that revisiting the epidemiologic transition theory in light of recent developments contributes to ongoing dialogue, thereby improving our ability to characterize demographic trends.

Our paper will focus largely on the U.S., but will also draw on relevant international research and data. In the first section of the paper, we will describe the epidemiologic transition theory as it relates to the demographic transition. This section will explore recent expansions of the theory which attempt to bring social, cultural, and behavioral factors to the fore. In the next section, we will offer a critique of the epidemiologic transition theory which discusses its limitations as a universalizing concept. This section will focus on the need to "particularize" epidemiological investigations and thus account for mortality and morbidity differentials between population subgroups. The final section of our paper will focus on the HIV/AIDS pandemic as both a case study and microcosm of the complexity of the epidemiologic transition. The AIDS pandemic will be used to demonstrate the continued presence and threat of infectious disease. In addition, the disproportionate impact of AIDS, due to social, cultural, economic, and behavioral factors, on particular population subgroups threatens to exacerbate mortality differentials (and other demographic trends) and thus seriously questions the universalizing tendency of the epidemiologic transition theory. Finally, we will attempt to demonstrate how the modern state of epidemic affairs, particularly AIDS and its related epidemics, transforms a current view of the epidemiological transition and more accurately accounts for the diversity of that transition for different subgroups of the population.

Epidemiologic Transition Theory

One of the central tenets of demographic theory is the concept of the demographic transition. The demographic transition refers to declining mortality and fertility that occurs in societies as they develop. Recognizing that this concept is broad, demographers have attempted to define it more precisely. One enhancement has been the development of the theory of epidemiologic transition. Omran first described this theory in 1971 (Omran, 1971) and later applied it to examine disease and mortality trends in the U.S. (Omran, 1977).

For purposes of explication, we shall first define epidemiologic transition theory in the context of Western (European) experience. The theory "focuses on the shifting web of health and disease patterns in population groups and their links with several demographic social, economic, ecologic, and biologic changes" (Omran, 1977). Olshansky and Ault (1986) provide a nice overview of the basic concepts underlying the epidemiologic transition theory:

As nations modernize they tend to improve their social, economic, and health conditions. Life conditions that were previously conducive to the spread of infectious and parasitic disease are rapidly replaced by more sanitary living conditions, improved medical technology, and better lifestyles. As the risk of dying from infectious diseases is reduced for a population, those saved from dying from such diseases survive into middle and older ages where they face the risk of dying from (chronic) diseases.

More specifically, epidemiologic transition theory embodies three primary stages through which mortality and disease patterns shift. The first stage is termed the "age of pestilence and famine". In the context of European experience, this first stage is "an extension of the pattern in the middle ages...and continued in Europe through most of the 18th century" (Omran, 1977). The first stage is marked by high mortality and fertility, and high prevalence of infectious diseases from which the high mortality derives. In this first stage, "mortality is high and fluctuating, thus precluding sustained population growth".

Gradually, the first stage gives way to the second stage, which is termed the "age of receding pandemics". During this time, the prevalence of infectious diseases begins to decline, and with it, mortality declines also. Furthermore, according to Omran, the rate of mortality decline "accelerates as epidemic peaks become less frequent or disappear", and societies begin to experience exponentially increasing populations. In the European context, this occurred during the 19th and early 20th centuries.

The third stage in the epidemiologic transition theory is the "age of degenerative and man-made diseases". In this stage, infectious diseases become increasingly rare, and they are replaced by chronic, degenerative, and stress-related diseases as the primary causes of death. According to Omran, in the third stage, mortality decline decelerates and stabilizes, life expectancy is high, and "fertility becomes the crucial factor in population growth".

Revisiting Omran's theory in 1986, Olshansky and Ault (1986) postulated that there is also a fourth stage of the epidemiologic transition. At the time Omran developed the theory, it was believed that declines in mortality from degenerative diseases would proceed at a very slow pace. However, life expectancies continued to improve in those countries that were experiencing this "final" stage in the transition. Thus, Olshansky and Ault, noting that what was occurring was a "postponement of the ages at which degenerative diseases tend to kill", suggested that the fourth stage should be characterized as the "age of delayed degenerative diseases".

The model outlined above is the classic or Western model. For societies following this model,
the transition is mostly complete, having occurred over the past 200 hundred years. Death rates have declined from roughly 30 per 1000 population to 10 per 1000, while fertility has declined from 35 births per 1000 population to 20 per 1000. These constitute changes from high rates to low rates. Omran characterizes these declines as "gradual" and occurring "in response to social, economic, and environmental improvements" that accompany modernization (Omran, 1977). Initial mortality declines are thus more attributable to environmental improvements (sanitation, etc.), while later declines are more attributable to improvements in medical technology and health care system development (aspects of the final steps in the completion of the transition through to the third stage).

Omran articulates two basic variants of the classic epidemiologic transition. One variant is the "accelerated model" in which the mortality transition was also socially determined, but was enhanced more powerfully by medical technology. According to Omran, this model applies to Japan, Eastern Europe, and the former U.S.S.R. Omran terms the second variant the delayed model, which applies to most of the Third World countries. In this model, mortality decline has only recently started to decline, although the rate of decline has been dramatic. The primary reason is large-scale importation of medical technologies and public health measures. Under this model, fertility declines have not equalled mortality declines, largely due to the rapid pace with which the mortality decline ensued. Omran notes that countries which have engaged in active population control programs represent a subcategory of this model.

A main corollary to the epidemiologic transition theory is that demographic changes both shape and are shaped by epidemiologic changes. Thus, since infectious diseases tend to impact the young most heavily, as a population shifts along the epidemiologic transition between the first and second stages, survival gains will be highest among children and women.* Increased numbers of surviving children produce a "wave of children and youths moving up through the population pyramid", while increased survival for women relative to men implies a shift in the population sex ratio. Then, as the population moves into the third (and fourth) stage, increases in life expectancy translate into an aging population. Omran also notes some important social implications which accompany the epidemiologic transition:

> Declines in morbidity and mortality can increase the efficiency and productivity of the labor force: healthier adults can work better and more surviving children means more potential workers....Improved childhood survival presumably removes the complex social, emotional and economic rationales for high birth rates (Omran, 1977).

Omran demonstrates the applicability of his theory by examining morbidity and mortality trends in the United States during the 19th and 20th centuries. Specifically, he examines overall U.S. data from 1890 through 1970, supplemented with data from Massachusetts and New York City from 1800 through 1970. He points to declining death rates, particularly among infants and children, shifts in leading causes of death from infectious to chronic diseases, and decreasing fertility. Furthermore, he discusses declines in maternal mortality due largely to better nutrition, better overall medical care, and particularly, better obstetrical care, especially during labor and delivery. Finally, he notes that overall mortality declines during the late 19th century were attributable to improved sanitation, public health measures (e.g. immunization, isolation, and quarantine), and that later 20th century gains were realized from increasing organization and quality of the medical care sector.

Omran also spends some time developing the above descriptions of aspects of mortality and morbidity change in the U.S. that occurred during the country's epidemiologic transition. Importantly, he discusses differential mortality patterns by subgroups. Thus, he notes that the transition favors the young over the old (because the declines occur as a result of reductions in infectious diseases, which exact their greatest toll on infants and children). Furthermore, according to Omran, the transition favors females over males:

> The female [of the 19th century and before] usually filled the roles of wife, mother, nurse, cook, and maid... She was overworked undernourished and frequently exposed to infection from several members of the family. With strains of repeated pregnancies, lengthy breastfeeding and prolonged childrearing, she was vulnerable to many illnesses, especially tuberculosis, anemia, and disease related to pregnancy and labor. When the infectious disease subsided and women began to have fewer pregnancies, female mortality decreased correspondingly (Omran, 1977).

Omran also notes that the transition favors whites over non-whites. He discusses that whites in U.S. society have always had advantages in terms of socioeconomic status, education, nutrition, and access to health care. Therefore, whites have correspondingly lower mortality and morbidity. Thus, for whites the epidemiologic transition "started earlier and moved faster. In particular the shift from communicable to degenerative disease has been slower among nonwhites...[who] still have higher death rates from infectious diseases...."

After noting differential mortality patterns in the U.S. by various subgroups and discussing some weaknesses with early U.S. mortality and morbidity data, Omran recapitulates epidemiologic transition theory, and then proceeds to return to the general
conclusions that the theory predicts. Thus, emphasizing the overall changes in mortality, morbidity, and fertility that occurred in the U.S. from 1800 to 1970, Omran concludes his paper: “data from the US...leave no doubt that the transition in the US in the last two centuries belongs to the Classical or Western Model of the epidemiologic transition.”

THE EPIDEMIOLOGIC TRANSITION REVISITED: THE HEALTH TRANSITION

As noted above, Omran briefly discusses subgroup differences in mortality, but these are secondary points to his theory. His primary emphasis is on the changes in mortality brought about by the shift from mortality patterns dominated by infectious diseases to patterns dominated by degenerative diseases. Differential mortality, then, becomes more a function of which groups are more prone to dying and sickness from infectious diseases, and thus stand to benefit most as the society undergoes the epidemiologic transition and infectious diseases are virtually eliminated. By the time Omran first published his theory, it was already the subject of some debate, and this debate has continued into the present. Recently, a prominent group of scientists convened an international conference to re-examine epidemiologic transition theory. The conference’s proceedings are titled What We Know About the Health Transition: The Cultural, Social, and Behavioral Determinants of Health (Caldwell et al., 1990). Two important emphases of the conference are evident from this title. First, the concept of the epidemiologic transition has been broadened to that of a “health transition”. John Caldwell (1990), in his introduction, notes:

A broader term than mortality transition has been the epidemiologic transition because it embraces changes in levels of sickness as well as mortality. For our purposes, neither term is sufficient because both are purely outcome measures. We want a term that includes the social and behavioral changes which parallels the epidemiologic transition....We employ the term health transition to include both epidemiological and related social changes.

Thus, the health transition theory does not relegated social changes to a secondary position, but includes them as a central focus, along with the basic outcome measures of mortality and morbidity which had been the focus of epidemiologic transition theory. Herein lies the second point evident from the title of the conference: scientists are going beyond the broad generalizations of the epidemiologic transition theory to focus in greater depth on other important factors (cultural, social, and behavioral) which influence health outcomes.

Lado Ruzicka and Penny Kane (Ruzicka and Kane, 1990) provide an overview of morbidity and mortality trends in the health transition. Their paper’s discussion of “the inequality of death” is of crucial importance for this analysis. They provide a thorough analysis of a fundamental truism in demography—that mortality patterns vary widely by race, sex, economic indicators, and class. Furthermore, the authors point out that there is substantial “heterogeneity within social class and, for that matter, within any other sociocultural, demographic or economic category” in terms of mortality risks. They also note an overconfidence embodied by the epidemiologic transition theory that societies will gradually progress to the point where they have virtually eliminated infectious diseases as a major health threat. In fact, they state, “the only disease that was eradicated and has not reappeared in more than a decade is small pox” and they conclude “[t]he emergence of AIDS...is a sobering warning that new disease may still emerge”.

Stephen Kunitz’s (Kunitz, 1990) discussion of the importance of sociocultural determinants of mortality differentials among subgroups is perhaps the most eloquent and comprehensive within this body of literature. Kunitz notes the value of the generalizations inherent in epidemiologic transition theory, but is quick to point out their problems:

The generalizations implied in these typologies are of course useful. It is possible to commit the historicist fallacy, however, and assume that development stages are everywhere the same and follow one another in some inevitable progression...while it is true that stages of development may be usefully regarded as ideal types—that is to say as models or heuristic devices—such constructs may be misleading because they may cause us to generalize inappropriately and to reify what began simply as an abstraction...(Kunitz, 1990).

Thus, Kunitz’s main issue is that overgeneralization of the theory is fraught with pitfalls. Kunitz warns that identifying and understanding important social and cultural differences among subgroups are essential components to making good health policy.

Kunitz notes several areas in which standard thinking about health improvements is useless when faced with powerful sociocultural forces that work against standard solutions. He discusses the common belief among health professionals that specific medical interventions can be mechanically applied to cure certain diseases. Kunitz points out that this thinking works in the case of a disease condition which can be attributed to a single primary cause, but falls short in the case of a disease condition which is best explained by “multiple weakly sufficient causes, or risk factors”. Indeed, such explanations are fundamentally important for understanding mortality differentials among subgroups in the U.S.

Two main points regarding limitations of epidemiologic transition theory emerge from the above discussion. First, large differentials in mortality trends among various population subgroups may undermine the generalizability of epidemiologic transition theory. Second, the epidemiologic tran-
Refocusing the lens

Expanding the discussion: Limitations of epidemiologic transition theory

As noted above, the two primary areas in which epidemiologic transition theory falls short of the mark are in its failure to particularize to important population subgroups, and in its suggestion that infectious diseases are gradually eliminated as a major health threat. We will first explore subgroup differences, focusing predominantly on the U.S., but also drawing on some international examples.

Blacks have consistently had higher mortality rates than whites throughout U.S. history. Moreover, as noted above, blacks have had, and continue to have, higher death rates from infectious diseases.* For example, age-adjusted influenza and pneumonia death rates for black males in 1986 were 2.72 deaths per 1000 population—white males were experiencing this same rate in 1950. The 1986 rate for white males is 1.75 deaths per 1000 population, or 35% lower than the rate for black males. One claim made by the health transition panelists is that while “total mortality differentials in the most advanced countries still persist...the onset of the health transition is marked by convergence [of mortality differentials], not divergence” (Palloni, 1990). In terms of blacks in the U.S., this claim appears to be incorrect. Figure 1 presents total age-adjusted death rates by race in the U.S. for the years 1900 through 1990. The death rates are shown on a logarithmic scale so that the differential can be more clearly seen. Convergence does not occur, as can be seen by the dashed line which indicates excess black mortality as a percent of white mortality—though mortality rates for both races declined during this period, race differentials do not show improvement: in 1900 black mortality rates were 58% higher than

*It should be noted that extensive literature searches found no articles which examined long-term trends in differential survival by subgroups, by year, and by cause of death. Omran’s piece and other articles in the literature (e.g., U.S. Department of Health and Human Resources, 1990) note that black declines in infectious diseases have been less than those of whites. Nevertheless, our search did not discover adequate data to effectively illustrate this trend. This lack of data and dearth of literature further emphasize the point that to date, subgroup mortality differentials, particularly in the context of the epidemiologic transition, have received inadequate attention.
white mortality rates; in 1990 the differential was 62%. Over the course of the 20th century, the differential exhibited marked fluctuation. It is worth noting that the publication of Omran’s theory (Omran, 1971) coincided with the minimum black–white mortality differential (the trough in the dashed line in Fig. 1). Since that time, the differential has been increasing.

Geographic areas often exhibit racial mortality differentials in their starkest light. Massey and Denton (1989) show that U.S. inner city areas display marked “hypersegregation”. In these areas, blacks are severely isolated from mainstream society across multiple social dimensions and indicators of this isolation are readily found, notably differences in fertility, language, employment, and family patterns. Massey and Denton do not explicitly point to health differentials in the inner city, but their analysis ties directly to recent work on this very topic. Two Harlem physicians analyzed excess mortality in Harlem and demonstrated that males in Harlem are less likely to reach age 65 than males in Bangladesh, and that the mortality rates in Harlem are more than twice those of U.S. whites. The researchers conclude that “Harlem and probably other inner city areas with largely black populations have extremely high mortality rates that justify special consideration analogous to that given to natural disaster areas” (McCord and Freeman, 1990). Even more relevant to the present analysis, a working paper from the Center for Population and Family Health at Columbia University describes the apparent reversal of the epidemiologic transition in Harlem:

...by 1990, AIDS, homicide, pneumonia, influenza, and other infections joined the ranks of major killers in this community. Thus, [a] major source of death in Harlem [is] early transitional infectious and respiratory diseases (Findley and Ford, 1993).

A recent article published by National Center for Health Statistics researchers in the Statistical Bulletin (Queen et al., 1994) discusses socioeconomic disparities in mortality and mortality trends using data from the 1986 National Mortality Followback Survey and the 1986 Health Interview Survey. In addition to noting the race differences discussed above (“the gap between the death rates of whites and blacks has widened even as the overall US death rate has declined”), the paper also concludes that disparities in mortality rates on the basis of income and education have increased over the years 1960 to 1986. For example, age-adjusted mortality rates among white males with high educational attainment decreased by 50% during this time period (from 5.7 to 2.8 deaths per 1000 population), while for white males with low educational attainment, the decrease was only 15% (from 9 to 7.6 deaths per 1000 population).

In a more complex, regression-based analysis of the same subject, Paul Menchik (1993) suggests that income differences are of primary importance in determining mortality differences between older black and white men over the period 1966–1981. The article concludes, differential mortality by economic status is strongly present in the United States today, and...this relationship is monotonic...differences in mortality between ethnic groups are, in large part, a consequence of poverty...as opposed to genotype (Menchik, 1993).

Numerous international evidence lends further credence to the thesis of this paper—that the epidemiologic transition is much less uniform and consistent a process than widely held notions would suggest. Ruzicka and Kane’s article (Ruzicka and Kane, 1990), in addition to showing point-in-time mortality differentials by income and race, contains tabulations of standardized mortality by social class in England for the year 1971. These data show virtually monotonic decreases in mortality as social class increases. The paper also presents standardized mortality ratios (SMRs) for 11 distinct international regions (including the U.S.) by social class (six classes) across the years 1971–1975. These SMRs allow for the comparison of mortality differentials between regions but within a social class.* The differences between regions for a given social class are striking, even when considering regions which are in the same general stages of economic development. For example, in India and Pakistan the standardized mortality ratio in the lowest social class was 122, while for the West Indies it was 267. The analysis also shows that social class mortality differentials between countries were attenuated at higher social classes—i.e., individuals of higher class do similarly well in any country.

Related international comparisons by John Caldwell suggest that “some countries reach health levels far above those that would be dictated by their economies and others fall far below” (Caldwell, 1986). The superior health achievers (e.g. Sri Lanka, Costa Rica) are characterized by high levels of female autonomy, emphasis on education, open political systems, and civilian societies without rigid class structures, while the lower health countries (e.g. Libya, Saudi Arabia) show a lack of these characteristics. While Caldwell’s approach is an oversimplification of a complex interaction of social phenomena, it provocatively encapsulates the essence of this interaction.

The above examples demonstrate that mortality experiences vary widely by different subgroups, both within the U.S. and across countries. The other general area to explore is the degree to which infectious diseases are truly being eliminated as societies develop. Again, substantial evidence suggests
that the epidemiologic transition theory's assertion that development leads to an eventual triumph over infectious diseases is overstated. Pinner et al. (1996) recently reviewed current trends in infectious disease mortality in the United States by studying ICD-9 codes and underlying cause of death for diseases such as tuberculosis, HIV, meningitis, and respiratory and GI infections. They conclude, "[d]espite historical predictions that infectious diseases would wane in the United States...infectious diseases mortality in the United States has been increasing in recent years". A recent article in the lay press reviewed (admittedly in a somewhat alarmist fashion) the upsurge in the incidence of epidemics in the U.S. (Lemonick, 1994). The article notes localized but non-trivial epidemics of whooping cough, lyme disease, malaria, streptococcus-A infections, hepatitis, and measles. It also points to the worldwide resurgence of cholera, particularly widespread outbreaks in Latin America.

In addition to AIDS (which will be discussed at length below), the accompanying resurgence of antibiotic-resistant tuberculosis represents the most concerning new epidemic. In fact, the resurgence of TB is largely due to the interaction between HIV infection and TB: "This relationship is synergistic...HIV multiplies the problems of tuberculosis for individuals and entire communities; tuberculosis complicates the management and course of HIV infection" (Mann et al., 1992). The Centers for Disease Control and Prevention (CDC) recently noted that in the U.S. tuberculosis increased by 26.5% from 1986 to 1989 among young black and Hispanic adults as compared to less than 0.5% for non-Hispanic whites (Centers for Disease Control, 1991). Important here, once again, is the concept of subgroup differences. Not only are infectious diseases still problematic, but they exact a heavier toll on disadvantaged groups. The recent measles epidemic in immigrant neighborhoods in Los Angeles also demonstrates this point. In terms of the broad categories of the epidemiologic transition, a recent study examining gender differences in the U.S. concluded, "infectious disease mortality declines more in males, while degenerative disease mortality declines more in females" (Gage, 1994). Finally, Rudolfo Bulatao (1993), in a 1993 National Research Council conference on the epidemiologic transition in developing countries, demonstrated that while developing countries exhibit declining trends in infectious diseases that are consistent with the epidemiologic transition, alternative projections of these trends through the year 2015 give widely varying estimates of where various countries will be in terms of distribution of causes of death.

Several prominent works have recently been published on the subject of emerging and persistent infectious diseases. Each book contains similar conclusions: "diseases long thought to have been defeated could return" (Garrett, 1994), and "many of these diseases may be prevented...and, in a very few cases eradicated—but the majority are likely to persevere. We can also be confident that new diseases will emerge..." (Lederberg et al., 1992). The research presented in these books highlights the interplay of opportunistic infectious diseases and major human events: "social and economic change, changes in human behavior, and catastrophic events such as war and famine...may fan a minor outbreak into a widespread epidemic" (Morse, 1993; Krause, 1981). Moreover, the books highlight the concern that human manipulation of the environment can create higher mortality risks from infectious diseases: "changes at the micro level of the environment of any nation can affect life at the global, macro level" (Garrett, 1994). An excellent example of these points is the emergence of extremely robust, insecticide-resistant mosquitoes as a result of failed attempts to eliminate malaria using DDT spraying. Another example is numerous cholera and dysentery epidemics within Rwandan refugee camps.

While the re-emergence of infectious diseases in Harlem (discussed above) offers a localized picture of a reversal of the epidemiologic transition, the deteriorating health conditions in Russia present a much broader case of the same phenomenon. Newly released statistics show that deaths in 1994 from infectious disease are up 17.9% from 1993, and that the incidence of several diseases (typhus, diphtheria, and measles) is up over 300% (Feshbach, 1995). The ephemeral nature of health trends in developed countries seriously undermines widely held notions of epidemiologic transition as a stable march of progress.

The above examples are not intended to provide exhaustive proof of the inadequacy of epidemiologic transition theory, nor to suggest that the theory is useless as a tool for understanding the relationship between mortality change and development. Rather, the examples serve to briefly illustrate some of the problems inherent in epidemiologic transition and health transition theory in order to constructively expand the current discussion so that it addresses these theoretical limitations. Generalizations of the theory mask important subgroup differences. These differences imply that optimism regarding death rates is unjustifiable in the case of certain subgroups, who risk being lost in the broader trends toward improved health shown by the population at large. Moreover, evidence abounds which seriously questions the notion that infectious diseases have been mostly conquered by modern science. In order to develop this point more fully, we will examine one case in depth—the HIV/AIDS pandemic.
THE HIV/AIDS PANDEMIC AND THE EPIDEMIOLOGIC TRANSITION

...[The AIDS pandemic has outgrown—both practically and intellectually—the capacities of the global public health system. It has also outgrown many of the conventional wisdoms about its importance and implications (Hamilton, 1994).

As of June 1994, over 400,000 cases of Acquired Immune Deficiency Syndrome (AIDS) had been reported to the CDC (Centers for Disease Control, 1994). The CDC estimates that there are over 1 million HIV infected persons in the United States and over 13 million HIV infected persons worldwide. Projections for the year 2000 range between 38 and 110 million HIV infected persons (Mann et al., 1992).*

AIDS, as a syndrome and collection of diseases and conditions, has taken particular geographic and population-specific shapes, resulting in a series of complicated and diverse epidemics throughout the world. Furthermore, AIDS is more than a medical phenomenon: the biological reality of HIV has intersected with the social realities of poverty, poor health status, homelessness, substance use, and discrimination to create disproportionate impacts on particular population subgroups (Panos Institute, 1992; Hamilton, 1994):

AIDS is socially constructed (as society perceives and frames the phenomenon, blames victims, and laboriously negotiates its response) yet at the same time fits nicely into a one-dimensionally reductionist and biologically based model of disease (Rosenberg, 1989).

AIDS represents both a unique challenge to an era which heralds the end of infectious diseases and epidemics and a reminder that history has demonstrated the persistent ability of disease to devastate populations, particularly the most vulnerable. As such, AIDS promises to both contribute to and challenge any theory of epidemiological transition. Furthermore, as an epidemic that may have considerable global economic and political consequences, it warrants attention from any theory which seeks to explain societal transitions of mortality and morbidity. In this section, we will outline some of the distinct characteristics of the AIDS epidemic, while recognizing its historical connections to past epidemics, and attempt to demonstrate how a consideration of the AIDS pandemic enhances and challenges the theory of epidemiologic transition in the United States.

It is important for any analysis of the AIDS epidemic to examine its distinct qualities as a modern day epidemic (Panos Institute, 1992; Hamilton, 1994).

1. AIDS is the first major epidemic to occur in the past 50 years, after the apparent conquering of infectious diseases;
2. AIDS has a case-fatality rate approaching 100%, with no cure currently available;†
3. AIDS is primarily transmitted through sexual contact, although there is an increasing percentage of cases due to injection drug use;
4. AIDS has an extraordinary capacity for growth, spreading rapidly into geographic areas and communities previously unexposed to the epidemic;
5. within each affected community, the epidemic has evolved and become more complex over time;
6. HIV has a long incubation period—an individual may remain asymptomatic for up to 10 years;
7. AIDS has prompted and/or exacerbated the spread of numerous other epidemics and diseases such as tuberculosis, other STDs, malaria and certain cancers; and
8. AIDS affects most individuals in their economically most productive and childbearing years.

At the same time, it must be remembered that AIDS is not wholly different from past epidemics, such as the bubonic plague and cholera. Garrett offers a reminder that “HIV is but one of a long series of microbes that have recently surfaced, and will be followed by more” (Garrett, 1992). As discussed previously, the emergence of full-fledged epidemics from incidental microbial infections is due to a series of factors that include globalization and economic interdependence, international travel, and interaction with socioeconomic conditions such as poverty and homelessness. The history of epidemics, of which AIDS is a part, is a clear indication that infectious diseases are a continual component of population change.

The demographic impact of AIDS

The distinct characteristics of the AIDS pandemic combine to produce significant demographic consequences in both developed and developing countries. The effects of AIDS can be seen at the micro (individual and family) and macro (societal/global) levels. Two of the most notable components of the demographic effect of AIDS are its apparent 100% case-fatality rate and its impact on individuals in their most economically productive and

---

*HIV projections vary largely due to different modeling techniques and assumptions. Mann et al. (1992), using the Delphi method, produce quite high estimates. The World Health Organization, on the other hand, uses much more conservative assumptions and low-end estimates. In general, however, there is no agreed upon projection technique, and the nature of the HIV disease itself, such as the long incubation period (up to 10 years) and the lack of individual knowledge of serostatus complicate measurement of incidence and prevalence.

†It is important to note that even if a cure for AIDS were discovered, global problems of access, poor health infrastructures, and cost would seriously impede its ability to end the AIDS epidemic.
reproductive years. At the family level, adults with AIDS require extensive household resources, particularly for health expenditures, at the same time that their ability to work increasingly deteriorates, and their contribution to household income decreases. The net effect of reduced household income impacts all members of the household, an impact which is greatly exacerbated when a parent or household adult dies from AIDS (Armstrong, 1992). Furthermore, "decisions made at the household level to reallocate resources...may, when aggregated play out at the sectoral and macroeconomic levels" (Hamilton, 1994). These demographic effects can most clearly be seen in the developing world and in populations of the developed world where AIDS is increasingly spread through heterosexual contact.

The demographic impact of AIDS on population growth largely depends on particular population characteristics and population diversity. For instance, high infection rates for women in their reproductive years could more seriously affect population growth and structure than similar rates for women past their reproductive years or for men. Surveillance studies and AIDS projections indicate that women face increasing rates of HIV infection, prompting the World Bank to estimate that AIDS may have a sizeable effect on population growth, one that will be increasingly evident over time:

The reason is the multiplier effect of mortality of adults of reproductive age: the deficit of people [initially] is almost exclusively due to AIDS mortality, which is only a small proportion of all deaths. Two or three decades from now, the reduced number of people will be due both to AIDS mortality and to births that did not occur because of the additional mortality among women of reproductive age (Armstrong, 1992).

Still another mortality effect is caused by the increasing number of infants born with HIV infection; perinatal transmission rates are estimated to be between 7 and 42%, with higher rates in the developing world (Mann et al., 1992).

AIDS has its most transparent demographic effects on mortality and morbidity. Recent projections using New York City data for males and females indicate increases in age-specific mortality rates due to HIV-related mortality (Kates and Weinstein, in press). In addition, AIDS significantly affects the morbidity of a population since it encompasses numerous diseases and conditions which result in increasing disability. Projections of life expectancy at birth are also affected by AIDS: first, AIDS may reverse previous trends of steadily increasing life expectancy, and second, in areas with hard-hit female populations, AIDS may reduce the gap in life expectancy which currently favors women over men. The Statistical Bulletin recently reported that preliminary life tables prepared by the Metropolitan Life Insurance Company indicate the first stagnation in improvements in U.S. life expectancy since 1980. Researchers attribute this to the relatively large increase in overall mortality, particularly due to AIDS and related diseases (Kranzzer, 1995). In addition to potential declines in life expectancy due to AIDS, positive trends in infant and child mortality may also diminish, particularly in areas where women have a high rate of HIV infection.

Another significant aspect of the AIDS pandemic is its interaction with other diseases and its capacity to create and/or exacerbate other epidemics:

AIDS is a unique disease; no other known infectious disease causes its damage through a direct attack on the human immune system. Because the immune system is the final mediator of human host–infectious agent interactions, it was anticipated early on that HIV infection would complicate the course of other important human diseases (Mann et al., 1992).

HIV infection has been associated with increased occurrence of cancers such as Kaposi's sarcoma, non-Hodgkin's lymphoma and, for women, cervical cancer. HIV has also been associated with increased incidence of other sexually transmitted diseases, a relationship which is both highly dynamic and synergistic (Berezin, 1992). In particular, there has been a correlation between HIV infection, gonorrhea, syphilis, chlamydia, human papillomavirus, and hepatitis B. In addition, it is important to note morbidity differentials during the course of disease progression. For instance, as women with HIV infection become increasingly immunosuppressed, they are more likely to develop gynecological complications than HIV negative women.

As discussed above, one of HIV's most serious effects, in terms of its demographic impact, has been its interaction with tuberculosis. This relationship is bi-directional—HIV further complicates TB treatment, while TB hampers care delivery of HIV infected persons. This synergism is particularly dangerous in developing countries where 95% of people with dual TB and HIV live, and in hard-hit urban centers of the developed world, where, until the advent of AIDS, TB was considered a "stable, endemic health problem". Estimates of the interaction demonstrate that in 1991, of the 1.72 billion TB infections and the 11.8 million HIV infections, 4.6 million individuals were TB and HIV infected (Mann et al., 1992).

In the U.S., the new demographics of TB show heavy concentrations of the disease in the most populous states. Like HIV, TB has historically been linked to environmental factors, such as poverty, crowding, and access to health care. The interaction of the two epidemics has been the most severe in areas with high rates of HIV infection, such as New York, California, Florida, Texas and New Jersey. These states rank as the top five in reported AIDS cases and, together, represent more than half of new U.S. TB cases. In urban areas of the U.S., the increase in TB cases is concentrated among young...
adults, particularly among ethnic and racial minorities, homeless people, and injection drug users. Studies show that individuals with HIV disease are at increased risk of developing active TB disease, that people with AIDS have a higher prevalence of TB, and that TB is often one of the first opportunistic infections associated with HIV, which in part prompted the CDC to expand the surveillance definition of AIDS to include TB as an AIDS-defining condition. Finally, recent evidence suggests that an individual with HIV may be more likely to become infected with TB after contact with a contagious individual (Office of Technology Assessment, 1993).

The epidemiology of AIDS in the United States

Of the over 400,000 reported number of AIDS cases for adults/adolescents and children in the United States, 85% are concentrated in metropolitan areas with populations of 500,000 or more. The majority of reported adult cases are among men (87%), with the predominant mode of transmission being homosexual/bisexual contact (61%). Injection drug use accounts for 21% of reported cases among men. Among women, who comprise 13% of cumulative reported adult AIDS cases, injection drug use and heterosexual transmission account for the majority of cases (48% and 36%, respectively). Blacks are disproportionately represented among reported adult cases for both men (29%) and women (54%) compared to a U.S. black population of 13%. This is also true for pediatric cases, of which 56% of reported cases are among blacks, reflecting the fact that 89% of all pediatric cases are due to mother with or at risk for HIV infection, and black women are also disproportionately represented among female AIDS cases (Centers for Disease Control, 1994).

Between 1987 and 1991, the age-adjusted death rate for HIV more than doubled, from 5.5 to 11.3 deaths per 100,000 population, moving HIV infection from 15th to eighth in the ranking of leading causes of death in 1992 (National Center for Health Statistics, 1994). In 1993, HIV became the leading cause of death for 25–44-year-olds in the United States (Centers for Disease Control, 1993b).

Changing trends of HIV infection and HIV mortality in the U.S. provide important epidemiological information about subgroup mortality differentials. In particular, mortality data indicate increasingly wide disparities based on sex and race, particularly where the two interact. The CDC reports that "racial and ethnic minority populations have been disproportionately affected by HIV infection and AIDS since the beginning of the epidemic in the United States" (Centers for Disease Control, 1993a). Through June 1994, half of reported AIDS cases were among blacks and Hispanics, although these two groups represent only one fifth of the total U.S. population (Centers for Disease Control, 1993a, 1994).

In 1991, the HIV infection age-adjusted death rate for black men (52.9 deaths per 100,000) was more than three times that for white men (16.7). The rate for black women (12.0) was more than nine times that for white women (1.3). In addition, HIV ranked much higher in cause of death rankings for both blacks and Hispanics as compared to whites (National Center for Health Statistics, 1994).

Women constitute the fastest growing group of newly diagnosed people with AIDS. Annual reported AIDS cases among U.S. adult/adolescent women increased 9% from 1991 to 1992, as compared to 2.5% for men (Centers for Disease Control, 1993b). The incidence of AIDS among women increased 20-fold between 1981 and 1990 in the U.S. (Melnick et al., 1994). The most recent update from the CDC reports that incidence of HIV is increasing more rapidly in women than men; women comprise 18% of new AIDS cases compared to 13% of cumulative cases (Centers for Disease Control, 1995a). In 1994 the male–female AIDS case ratio was 4.9:1 compared to 14:1:1 in 1984 (Centers for Disease Control, 1993b, 1994). Although black and Hispanic women constitute 21% of all U.S. women, they account for 74% of women diagnosed with AIDS in the U.S. and 77% of cases reported in 1994. Case rates for black and Hispanic women were, respectively, 1.6 and 7 times greater than for white women (Centers for Disease Control, 1995a). Finally, in some hard-hit urban areas, such as New York City, AIDS has been the leading cause of death among women age 25–44 since 1988 (New York City Office of Vital Statistics and Epidemiology, 1988–1990).

The most recent study of HIV infection in women (Melnick et al., 1994) found that the risk of death differed for men and women with AIDS. Women had higher death rates than men even though disease progression rates did not differ significantly. More women than men experienced death as a first event, an effect that was exacerbated by race and/or injection drug use. In addition, the study found that although overall disease progression rates were similar for women and men, the incidence of certain opportunistic diseases varied by sex. Women, for instance, were significantly more likely to develop bacterial pneumonia. The researchers conclude that because "death was the first disease progression event for more women than men...observed survival differences may reflect a differential access to or utilization of health care resources by gender".

AIDS and the epidemiologic transition theory

Omran’s theory of the epidemiologic transition was developed for its explanatory and descriptive applicability to the relationship between changing mortality patterns and the demographic transition; specifically, the decline of mortality rates in both developed and developing countries due to modern-
ization, a decline in infectious diseases, and a subsequent improvement in social, economic and health conditions.

Consistent with our earlier, more general discussion, the introduction of AIDS into the global health economy questions two components of the epidemiologic transition theory. First, AIDS challenges the theory's assumption of the disappearance of infectious diseases and their accompanying mortality concentration at younger ages. With the decline in infectious diseases, the age structures of most countries shifted to reflect an older age distribution, one which was much more susceptible to degenerative diseases. Yet AIDS is an infectious disease which largely affects the younger age groups (individuals in their most economically productive and childbearing years). AIDS may have considerable demographic consequences not only on some developing countries, but on many hard-hit urban areas of developed countries, areas which are afflicted with poverty and inadequate health services. Furthermore, evidence has shown that AIDS has exacerbated, and perhaps even caused in some areas, a resurgence in other epidemics, such as TB. Given the current disproportionate impact of HIV infection on minorities, women, the poor, and young adults, and global AIDS projections which indicate significant impacts on past demographic trends, the AIDS pandemic highlights the need for a reformulation of the epidemiologic transition discussions.

The second limitation of the epidemiologic transition theory is its failure to adequately address mortality differentials between subgroups. Although Omran recognized that the decline in mortality rates was not the same for all populations within the United States—namely that minorities lagged behind whites in mortality decline—his theory nevertheless attempted to broadly generalize the epidemiologic transition. Yet, as Kunitz (1990) reminds us, such generalization masks astonishing diversity and fails to include many explanatory variables that are of significance if we are adequately to understand the distribution and change of health and diseases in populations, particularly their cultural, social, and behavioral determinants.

As we have noted above, Kunitz's work demonstrates the importance of differences in mortality declines among cultures: "One must finely analyze the particularistic features of nations, cultures, regions, and localities". AIDS, as both a case study of a modern epidemic and a microcosm of the "inequality of death" (Ruzicka and Kane, 1989) seriously questions the comprehensive applicability of the epidemiologic transition theory to all population subgroups.

Epidemiologic information concerning AIDS presented above demonstrates the increasingly disproportionate impact of AIDS on particular subgroups of the U.S. population. This demographic impact has serious repercussions both for past trends in mortality decline and future projections of population growth, death rates, life expectancy and fertility. In fact, the effects of AIDS on particular subgroups of the U.S. population recall earlier stages of the epidemiologic transition theory. Just as the "age of pestilence and famine" (the first stage) is characterized by high mortality, high fertility, and the prevalence of infectious diseases, so is the current epidemiologic situation of U.S. blacks in urban areas hard-hit by AIDS and TB. These epidemics have increased black mortality at younger ages, an effect exacerbated by disproportionate poverty. Increasing death rates and mortality differentials for blacks due to AIDS threaten past trends in increasing life expectancy and an aging population (the third and fourth stages of the epidemiologic transition). Furthermore, previously lower mortality rates among women are diminishing as a "shift has occurred in the worldwide pattern of AIDS incidence from a disease primarily among men to a pattern of gender equity" (Melnick et al., 1994). The larger implications of this trend are substantial: "The HIV infection/AIDS problem in women and children will doubtless become one of the major challenges to public health, health care, and social support systems worldwide" (Chin, 1990). The double effects of race and gender further demonstrate our argument. For instance, as the leading cause of death for black women age 25–44 in New York since 1988, AIDS will have a serious effect on the mortality differential between white and black women (Centers for Disease Control, 1993b). In addition, AIDS will have a mortality multiplier effect due to births that did not occur due to AIDS mortality among black women of reproductive age and to the increasing number of children born with HIV. Furthermore, black women tend to be disproportionately more impoverished as compared to white women, therefore indicating an increasing demand for (less accessible) public health services, a severe drain on household resources and loss of income.

The effects of AIDS are likely to exacerbate the mortality differentials between whites and racial and ethnic minorities and negatively affect female life expectancy gains, as AIDS has moved up the rank order for cause of death in the U.S. This epidemiologic reality seriously questions the epidemiologic transition theory's assumption that mortality due to infectious disease continues to decline and that this can be applied broadly to whole populations. The disproportionate impact of AIDS, its effects on particular populations due to social, cultural and economic differences and biological realities will have both micro and macro level demographic effects that challenge the universal and optimistic picture presented by the epidemiologic transition theory. As Kunitz (1990) states,
Many health problems in both rich and poor countries are still best explained by weakly sufficient causes, or risk factors. Understanding their incidence, prevalence, and distribution, as well as their treatment and prevention, may require intimate understanding of particular people and settings.

The challenge before us, then, is to seek to develop such an understanding in order to effectively intervene and resolve health problems such as HIV/AIDS.

CONCLUSION

In this paper, we have attempted to demonstrate how the epidemiologic transition theory, designed to explain global trends in the dynamic relationship between epidemiology and demographic change, only partially serves to explain mortality declines over the last century. Instead, drawing on Kunitz (1990), we argue for the need to "particularize" the focus of epidemiologic transition on population subgroups. This approach provides a fuller and more accurate account of the epidemiologic transition, thereby enhancing the existing theoretical framework.

Specifically, we have shown that particularized attention to population subgroups within the U.S. demonstrates that all subgroups have not experienced the epidemiologic transition in the same way. In fact, contrary to the epidemiologic transition theory's claim that mortality differentials between U.S. blacks and whites converge over time, we have found that these differentials actually fluctuate and have most recently been diverging. Furthermore, we argue that the theory has overstated the decline in infectious diseases as cause of death. We focus on the AIDS pandemic as a case study which both argues that the theory has overstated the decline in mortality differentials between U.S. blacks and whites, and demonstrates how mortality differentials between population subgroups are actually increasing—with serious demographic effects.

The AIDS pandemic also challenges the theory's statement that women are favored over men in the epidemiologic transition, as AIDS shows how women, due to environmental and behavioral factors, are increasingly impacted by the epidemic.

We hope that the argument presented by this paper will serve to enhance the theory of the epidemiologic transition by arguing for a particularized lens. This technique results in a more complex and comprehensive picture of multiple epidemiologic transitions, and demonstrates how population subgroups experience these transitions differently. Finally, we hope that such an approach will be used to guide public policy endeavors that seek to address epidemics such as AIDS and the inequality of death that still prevails in the U.S. and throughout the world.

Acknowledgements—This work commenced while the authors were graduate students at Princeton University's Woodrow Wilson School of Public and International Affairs, and Office of Population Research. The opinions expressed are those of the authors and do not necessarily represent the opinions or policies of The Lewin Group or Princeton University. The authors thank Thomas Espenshade of the Office of Population Research, Princeton University, and Maxine Weinstein of Georgetown University for their valuable comments on previous drafts.

REFERENCES


Centers for Disease Control and Prevention (CDC) (1991) Tuberculosis Elimination—USA. CDC, Atlanta, GA.

Centers for Disease Control and Prevention (CDC) (1993a) Fact Sheet, HIV/AIDS and Race/Ethnicity. CDC, Atlanta, GA.

Centers for Disease Control and Prevention (CDC) (1993b) Fact Sheet, Women and HIV/AIDS. CDC, Atlanta, GA.


Melnick et al. (1994) Survival and disease progression according to gender of patients with HIV infection. *Journal of the American Medical Association* 272(24), 1915–1921.


