Social Epidemiology in South Africa

Landon Myer1,2, Rodney I. Ehrlich1, and Ezra S. Susser2,3

1 School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa.
2 Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY.
3 New York State Psychiatric Institute, New York, NY.

Received for publication October 9, 2003; accepted for publication December 23, 2003.

INTRODUCTION

In this paper, we review the history, present state, and future of social epidemiology in South Africa. In “developing” nations such as South Africa, social epidemiology is cast in a new light owing to these countries’ distinctive political, economic, and social histories. South Africa is still grappling with the public health legacy of the colonial and apartheid eras while contending with new public health threats that are linked to a changing global economy, as well as the devastating human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemic. Furthermore, the possibilities for research in social epidemiology have been limited by the scarcity of resources and political constraints. While such factors are common to many developing countries, they are thrown into sharp relief in the case of South Africa.

We begin by tracing a brief history of social epidemiology in South Africa, noting its remarkable early growth. Then we review intriguing, if sparse, evidence on the current interplay between society and health in South Africa. We suggest that the structural conditions of the past, most notably the racial discrimination and labor migration policies systematized by apartheid-era governments, underlie much of the current variation in population health in South Africa. However, we also point to evidence of a “bipolar” epidemiologic transition superimposed upon this legacy and interacting with it. Looking to the future, we propose that social epidemiology can and must play an essential role in meeting the public health challenge of the HIV/AIDS epidemic.

South Africa is a middle-income country of 45 million people with dramatic disparities in wealth, despite a gross domestic product per capita of US $8,900 (placing it alongside countries like Brazil and Mexico in terms of 1999 purchasing power) (1). In addition, racial terminology has a complex history in South Africa, having been used as a means of social division and control (2, 3). For the descriptive purposes of this review, we use the “population group” racial nomenclature formalized under apartheid, since most socioeconomic and health data have been collected and reported in this way for decades. The four categories are African (alternatively Black), comprising 79 percent of the 2001 population; Coloured, a widely heterogeneous group sometimes referred to as persons of mixed ancestry (9 percent of the population); Asian (alternatively Indian) (2 percent); and White (alternatively European) (10 percent). In some contexts, the first three groups have been referred to as Black (4).

HISTORY OF SOCIAL EPIDEMIOLOGY IN SOUTH AFRICA

Unlike most developing countries, South Africa has a long history of research into the social and economic determinants of health. From the 1940s onward, research that may be characterized as social epidemiology made important contributions to the understanding of health disparities under apartheid and, more generally, in societies characterized by inequality. In the long term, this understanding informed demands for improved population health as part of the movement toward social justice.

The social determinants of population health have been debated in South Africa since the colonial period. While the interest in “tropical medicine” that emerged in Europe and the United States during the late 1800s was focused largely on the health of colonists and imperial armies (5), a separate discourse in medical and scientific circles was concerned with whether the health disparities between European and native populations were attributable to innate properties,
physiologic characteristics, or the socioeconomic conditions under which the populations lived (6). In South Africa, the most stigmatized illnesses of the era, including syphilis and tuberculosis, featured prominently in such debates. In the case of syphilis, the argument centered on whether the increased occurrence of venereal disease among Africans was due to biologic predisposition, an innately lascivious lifestyle, or some other cause (7, 8). For tuberculosis, the debate focused on whether the persistently higher rate of disease observed among Africans was attributable to a lack of acquired immunity to the bacilli introduced by colonial settlers or the impoverished living conditions most Black South Africans faced (9–12).

Discussions of the causes of disease in indigenous African populations drew heavily from the then-dominant, and largely prejudiced, European views of race and health. The first health care administrators to advocate for a public health approach among the indigenous population in South Africa, including A. G. Park-Ross and H. S. Gear, worked against this view during the early 1900s in their investigations into the health of African populations, especially in the rural “native reserves” of Zululand and the Transkei (13–15). Subsequently, some early South African epidemiologists helped to provide a critique of the notion that disease within indigenous groups was attributable to innate predispositions by presenting data on the distribution and socioeconomic determinants of disease in African populations. Foremost among these was Sidney Kark, whose work during the 1940s and 1950s at the rural Pholela Health Centre in Natal built on the foundation laid by Park-Ross and others (16, 17). Kark’s work demonstrated that the socioeconomic conditions of rural communities were decisive in shaping their morbidity profile (18–20). In his study of the epidemiology of syphilis, Kark noted that the system of labor migration enforced by the colonial government acted as a “social pathology” in facilitating the spread of venereal disease among African populations in urban and rural areas (21). He also suggested that the same basic interventions that were afforded to White South Africans could be effective in improving health in Black communities. Rather than viewing tuberculosis as an inevitable consequence of underdeveloped immunity, Kark followed Gear’s direction in promoting improvements in living conditions and nutrition to address tuberculosis in rural communities (22).

Kark’s work included the training of students in Natal who would later make important contributions to the development of social epidemiology. After working with Kark at Pholela, John Cassel went on to play a central role in the development of social epidemiology in the United States in the 1960s and 1970s (22, 23). Another student of Kark’s, Guy Steuart, became a leader in the field of health promotion, advocating for a discipline that extended beyond health education to recognition of the role of structural conditions in shaping individual health (24). Mervyn Susser studied with Kark at the Institute of Family Medicine in Durban in the late 1950s and later coauthored Sociology in Medicine with anthropologist Bill Watson. This book was among the first works to draw examples from both the developing and developed worlds to investigate how the social, cultural, and economic features of populations combine to influence patterns of health (25, 26). In addition, an American student of Kark’s, Jack Geiger, returned from his time in South Africa to play a leading role in community health care in the United States (27). Kark also trained scores of students after he left South Africa in the 1960s, and later had a profound influence on the development of the primary health care movement (28, 29).

Despite these promising origins, the potential of social epidemiology in South Africa was not fully realized during the years after 1960. This can be attributed in part to the repressive policies of the apartheid government, which led to the arrest of activists concerned about the socioeconomic and health impacts of apartheid, as well as the emigration of many progressive thinkers (including Kark). Nonetheless, the interrelations between changing socioeconomic conditions and the health status of populations remained a dominant issue in South African public health in the following decades. In the 1970s, papers by Cyril Wyndham and other investigators based at the parastatal Medical Research Council compared the mortality rates of different racial groups (30, 31). These investigations confirmed a social gradient in mortality rates, with a large disadvantage in all-cause mortality, particularly infant mortality, being observed among African and Coloured persons relative to Whites (with rates among Asians falling between). The causes of mortality also differed across racial groups. For example, in adults aged 15–64 years, analyses from this period ranked ischemic heart disease, motor vehicle accidents, stroke, and digestive cancers as the most common causes of death among Whites, as compared with “ill-defined causes,” homicide, stroke, and tuberculosis among Blacks (31). However, official cause-specific statistics were very crude, and because of the poor quality of death registration and certification, “ill-defined causes” remained a common category of official mortality among Blacks for many years (32). In addition, the data were skewed by the poor representation of periurban Black townships and by the exclusion of so-called homelands, the impoverished rural areas designated for Africans under apartheid.

These differences between racial groups were conventionally framed as indicative of the epidemiologic transition, with Whites representing the “First World” experience of mortality and morbidity and Blacks the “Third World” experience (30, 33). Secular socioeconomic change in the Black population (often characterized as urbanization or Westernization) was typically expected to accelerate the transition to “First World” patterns of chronic disease (26). For some authors within South Africa, however, these health disparities indicated the need for changes in health policy. In a seminal article, Wyndham and Les Irwig used the variations in health between racial groups to argue for the redistribution of health care budgets toward an emphasis on priority diseases, comprehensive public sector primary care, and poverty alleviation (30). However, publication of a companion piece by Ehrlich describing the socioeconomic disparities between racial groups was vetoed by the Medical Research Council, illustrating the boundary the conservative medical establishment sought to place between what it viewed as acceptable health research and that which entered the “political” arena (34).
Until the 1970s, occupation was relatively neglected as a focus of social epidemiology in South Africa, although the mining and industrial sectors, which formed the basis of wealth in the South African economy, reproduced the country’s dominant sociopolitical features. By the second half of the 20th century, production workers were mostly Black and were subject to greater hazardous occupational exposures than Whites. While political opposition to apartheid was stifled by repressive state measures implemented during the 1960s and 1970s, an emergent labor movement enjoyed relatively greater leeway in protesting poor working and living conditions. In these struggles, the movement typically emphasized class solidarity rather than racial solidarity. Concomitantly, a new generation of occupational epidemiologists began to investigate the epidemiology of common diseases such as tuberculosis and hypertension from an occupational-class rather than a racial-group perspective (35, 36). State and industry influences on the interpretation of epidemiologic information to the detriment of exposed and affected workers were also subject to critique by Myers in his review of the South African asbestos industry (37).

During the 1980s, epidemiologic methods were also directed toward demonstration of the health effects of political violence (38, 39). Using community-based surveys, Derek Yach measured the adverse effect of political unrest on the provision of amenities such as water and transportation, on utilization of health services for chronic conditions such as tuberculosis and hypertension, and on injury rates in affected township areas (40). Other commentaries from this period focused on unequal access to health services as a critical pathway through which apartheid policies shaped the health of different racial groups (41). Differences in health status between racial groups were related explicitly to the apartheid system by more politically inclined researchers who were members of the anti-apartheid National Medical and Dental Association (42, 43). These individuals argued that the World Health Organization objective of “Health For All” in South Africa required the elimination of apartheid at the political level as well as socioeconomic advances (43). Many of these activists contributed to the 1990 Maputo Conference, which is recognized as a seminal event in the transition from oppositional critique to planning for a new public health system in South Africa (44).

The interplay between social factors and health throughout the 20th century was usually discussed in terms of racial differences. However, there were notable examples of an increased disease burden’s being blamed on socioeconomic disadvantage within the White population. For example, driven in part by epidemics of tuberculosis and silicosis, White miners repeatedly protested against working and living conditions during the early 20th century (45). They won substantial improvements in remuneration, housing, and illness compensation, while their Black counterparts continued to endure poor working and living conditions (46).

CONTEMPORARY SOCIAL EPIDEMIOLOGY IN SOUTH AFRICA

Much of the patterning of health, poverty, and race observed in contemporary South African society is the result of the enduring effects of social, political, and economic discrimination. Extreme disparities in wealth and income persist; the national Gini coefficient for income inequality (0.58) demonstrates the second-highest level of inequality among all countries worldwide (47, 48). These disparities are generally aligned with race, and in most instances they make the separation of racial categories from socioeconomic status categories difficult (figure 1). Any investigation of disease etiology in this setting—not just research that could be defined as social epidemiology—must face the challenge of complex social gradients in health, measuring wealth and poverty within and across diverse rural and urban communities, interpreting the meaning of associations between race and health outcomes, and understanding the interactions between changing labor patterns, forms of livelihood, and population health.

Social gradients and health

In South Africa, as in most nations, social disadvantage is associated with increased morbidity and mortality (49, 50), but the relations between social stratification and health are far more complex than a simple, monotonic gradient. This is partly because South Africa is caught in what has been described as a “bipolar” epidemiologic transition (51, 52). The general shift of a middle-income country toward an increased burden of chronic diseases, such as ischemic heart disease, stroke, and cancer (53), has been joined by a rise in infectious diseases associated with HIV/AIDS and the persistence of the classic diseases of poverty, such as childhood malnutrition, gastroenteritis, tuberculosis, and pneumonia (54).

The interplay between this bipolar transition and social gradients in health is still not well understood. It is tempting to presume that chronic diseases will affect predominantly persons of higher socioeconomic status, while infectious diseases will involve primarily the poor. However, the relations appear to be far more complex than this. Comparing racial groups, infectious diseases are certainly most prominent in the relatively poor Black majority, but this does not imply an inverse association between infectious disease and the socioeconomic position of individuals within any racial group. In fact, in the case of HIV, there are examples suggesting that in some settings, relatively higher socioeconomic status may offer no protection against infection and may even have a detrimental effect (55–57). Similarly, with respect to chronic disease, findings do not suggest any simple correlation between increasing affluence and chronic disease within racial groups. The complexity is illustrated by the findings on hypertension in the recent national Demographic and Health Survey. The data showed very little association of hypertension with race, while its associations with wealth and urbanization appeared to be modified by gender (58–60) (see table 1; note that these data were not reported in age-standardized form, further clouding interpretation). In another example, as Wyndham documented in the 1970s (30), stroke is a prominent cause of death among both Blacks and Whites. The reasons for this pattern remain unclear; in fact, different forces may be driving stroke rates in these two groups, making the similarities deceptive (61). Such
complexities are unlikely be resolved without more sources of data.

One of the most striking relations in social epidemiology in South Africa pertains to injury. “Unnatural” death, including death due to violence, accident, or suicide, was the leading reported cause of mortality in the South African adult population between 1997 and 2001 (and is currently the second leading cause of death after HIV/AIDS) (54, 62, 63). Injury is neither an infectious disease nor a chronic disease, and it is not accounted for in the traditional descriptions of epidemiologic transition (64). Rates of both intentional and accidental injury are at least 1.5 times higher in urban communities than in rural communities (58), and there is evidence to suggest substantial racial (65) and socioeconomic (66, 67) variations in both the degree of violence and the forms of violence experienced (68). In addition, violence against women is a major component of violent crime, especially in poorer communities (69, 70). However, despite its significance for population health, there has been relatively little epidemiologic research on the etiology and health impact of violence in South Africa, including the effects on public mental health (71, 72).

The relations between health and poverty in South Africa are changing as the South African economy slowly grows and shifts focus (61, 73–77). Generally, the effects of rapid and ongoing urbanization on health are so distinct that some recent national health data have been presented with findings for the African majority separated into urban and nonurban categories (58). One of the major lifestyle changes associated with urbanization is smoking, the prevalence of which is increasing in the middle and lower socioeconomic strata, particularly in urban areas (78–80). This trend parallels findings elsewhere in the developing world and, as in other settings, will probably contribute to the increasing burden of chronic disease among the Black majority in the coming decades (81–83). In addition, with the widening access to middle-class living standards, some families from relatively poor backgrounds are rapidly becoming more affluent. This raises important questions about which social conditions—historical poverty, contemporary affluence, or the transition between them—are likely to have the greatest influence on these families’ overall health (84, 85).

Measurement of socioeconomic position

While the wide disparities that characterize South African society make gross associations between poverty and ill health unmistakable, finer analyses are rendered more challenging by the difficulty of measuring socioeconomic position with precision. Traditional indicators of socioeconomic position, such as social class or socioeconomic status, may be useful for demonstrating gross inequalities between rich and poor or for distinguishing relative positions within the highest socioeconomic stratum. However, use of these measures within poorer communities that represent the majority of the population can be problematic, given the variability in indicators of wealth in different settings (48, 86). Income and expenditure, whether measured at the individual level or at the household level, are often inadequate proxy measures owing to the existence of extensive family- and community-based networks for sharing resources (87, 88). The use of education as a measure of socioeconomic

![Figure 1](https://example.com/figure1.png)

**FIGURE 1.** Gross monthly incomes of economically active South Africans (persons aged 15–65 years who were either employed or seeking employment) by racial group, 2001. At the current exchange rate, 1 Rand (R) = US $7. Data were obtained from Statistics South Africa (4).
status can also be problematic, most notably in rural areas, where the historical underdevelopment of schools means that relatively few older adults have received postprimary schooling and therefore adult socioeconomic position often correlates poorly with level of formal education (89). Following international standards, South African researchers have developed a range of more broadly applicable measures of socioeconomic position for public health research. Although the deployment of different measures is often highly context-specific, with certain measures of socioeconomic position being preferred in certain settings, there has been some success in developing measures that can be applied across the country’s diversity. For example, the Asset Index, initially developed for the analysis of Demographic and Health Survey results from different countries, has been successfully adapted to South African data sets (90). This index is comprised of a series of measures assessing relative wealth or poverty (table 2) and represents a valuable approach to the measurement of socioeconomic position across diverse urban and rural settings.

The links between geographic area and health, a source of great interest in contemporary epidemiology (91), are marked in South Africa because of the residential segregation of African, Asian, Coloured, and White populations enforced during the apartheid era. As a result, area-level measures of socioeconomic context are of considerable interest and are becoming increasingly common in South African health research (92, 93). For example, area measures incorporating local levels of income, unemployment, welfare, and residential overcrowding have been applied to the study of associations between social disadvantage and adolescent asthma. The results suggested that while increased wealth may be associated with an increased prevalence of asthma, living in a poor community is associated with more severe symptoms of disease (94).

Residential location in South Africa is likely to represent more than just poverty, since the former townships where most Black South Africans live represent a radically different social context than the suburbs inhabited by most Whites, with areas historically reserved for Asian and

### TABLE 1. Relative odds of prevalent hypertension* in a representative sample of South African men and women, 1998†

<table>
<thead>
<tr>
<th></th>
<th>Men (704 hypertensive, 5,049 normotensive)</th>
<th>Women (1,280 hypertensive, 6,793 normotensive)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR‡ 95% CI‡</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Asset Index (in quintiles)§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorest fifth</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Second-poorest fifth</td>
<td>1.14 0.78, 1.68</td>
<td>1.00 0.74, 1.35</td>
</tr>
<tr>
<td>Middle fifth</td>
<td>1.11 0.75, 1.67</td>
<td>1.00 0.73, 1.36</td>
</tr>
<tr>
<td>Second-wealthiest fifth</td>
<td>1.77 1.15, 2.72</td>
<td>1.33 0.94, 1.87</td>
</tr>
<tr>
<td>Wealthiest fifth</td>
<td>2.20 1.35, 3.59</td>
<td>1.13 0.75, 1.72</td>
</tr>
<tr>
<td>Years of education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1–7</td>
<td>0.99 0.70, 1.39</td>
<td>1.19 0.94, 1.50</td>
</tr>
<tr>
<td>8–12</td>
<td>1.19 0.84, 1.68</td>
<td>1.11 0.84, 1.46</td>
</tr>
<tr>
<td>&gt;12</td>
<td>0.75 0.43, 1.31</td>
<td>0.58 0.36, 0.92</td>
</tr>
<tr>
<td>Racial group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Coloured</td>
<td>0.87 0.61, 1.25</td>
<td>1.40 1.04, 1.89</td>
</tr>
<tr>
<td>White</td>
<td>1.28 0.84, 1.96</td>
<td>1.23 0.87, 1.76</td>
</tr>
<tr>
<td>Asian</td>
<td>1.02 0.62, 1.70</td>
<td>1.08 0.71, 1.65</td>
</tr>
<tr>
<td>Area of residence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Rural</td>
<td>1.06 0.82, 1.37</td>
<td>0.63 0.50, 0.80</td>
</tr>
</tbody>
</table>

* Hypertension was defined as blood pressure ≥160/95 mmHg and/or current use of antihypertensive medication.
† Data were obtained from the 1998 Demographic and Health Survey (90, 165).
‡ OR, odds ratio; CI, confidence interval.
§ The Asset Index (90) is a summary measure of socioeconomic position based on a nine-item scale: source of drinking water, type of toilet facility, fuel used for cooking/heating, number of rooms in the home used for sleeping, affordability of food, and ownership of specific assets (also see table 2).
TABLE 2. Distribution of the poorest 40% and the wealthiest 20% of respondents to the 1998 Demographic and Health Survey according to the domains used to construct the Asset Index* for measurement of socioeconomic position, South Africa, 1998

<table>
<thead>
<tr>
<th>Domain and possible responses</th>
<th>Distribution (%)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poorest 40%</td>
</tr>
<tr>
<td>Source of drinking water</td>
<td></td>
</tr>
<tr>
<td>Piped water in home</td>
<td>3</td>
</tr>
<tr>
<td>Piped water on site, outside home</td>
<td>16</td>
</tr>
<tr>
<td>Piped water from public tap</td>
<td>40</td>
</tr>
<tr>
<td>Surface water from dam, spring, or stream</td>
<td>29</td>
</tr>
<tr>
<td>Type of toilet facility available</td>
<td></td>
</tr>
<tr>
<td>Flush toilet</td>
<td>3</td>
</tr>
<tr>
<td>Pit latrine</td>
<td>55</td>
</tr>
<tr>
<td>No toilet facility</td>
<td>29</td>
</tr>
<tr>
<td>Fuel used for cooking and/or heating</td>
<td></td>
</tr>
<tr>
<td>Electricity</td>
<td>4</td>
</tr>
<tr>
<td>Gas</td>
<td>4</td>
</tr>
<tr>
<td>Paraffin/kerosene</td>
<td>55</td>
</tr>
<tr>
<td>Wood</td>
<td>58</td>
</tr>
<tr>
<td>Mean no. of household members per room used for sleeping in the home</td>
<td>2.47</td>
</tr>
<tr>
<td>Main flooring material in the home</td>
<td></td>
</tr>
<tr>
<td>Earth</td>
<td>40</td>
</tr>
<tr>
<td>Cement</td>
<td>42</td>
</tr>
<tr>
<td>Vinyl</td>
<td>11</td>
</tr>
<tr>
<td>Carpet</td>
<td>5</td>
</tr>
<tr>
<td>Main wall material in the home</td>
<td></td>
</tr>
<tr>
<td>Mud</td>
<td>37</td>
</tr>
<tr>
<td>Mud and cement</td>
<td>16</td>
</tr>
<tr>
<td>Corrugated iron</td>
<td>13</td>
</tr>
<tr>
<td>Plaster</td>
<td>15</td>
</tr>
<tr>
<td>Affordability of food</td>
<td></td>
</tr>
<tr>
<td>Household often goes hungry</td>
<td>19</td>
</tr>
<tr>
<td>Household sometimes goes hungry</td>
<td>50</td>
</tr>
<tr>
<td>Household never goes hungry</td>
<td>25</td>
</tr>
<tr>
<td>Household and individual ownership of specific assets</td>
<td></td>
</tr>
<tr>
<td>Electricity in home</td>
<td>18</td>
</tr>
<tr>
<td>Sheep/cattle</td>
<td>20</td>
</tr>
<tr>
<td>Automobile</td>
<td>3</td>
</tr>
<tr>
<td>Telephone</td>
<td>1</td>
</tr>
<tr>
<td>Television</td>
<td>18</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>6</td>
</tr>
</tbody>
</table>

* Source: Bradshaw and Steyn (90).
† Distributions are presented as the percentage of households meeting the specified criterion, with the exception of mean number of household members per room used for sleeping. Only the most common responses are shown; as a result, percentages for some domains may total less than 100.

Coloured populations likely to fall somewhere in between. Even for better-off residents of African, Coloured, and Asian areas, factors such as violence (66, 72, 95), environmental exposures (96, 97), psychological stress (98–100), and the limitations of local medical services (76, 101) are likely to affect health, despite these residents’ relatively higher incomes in comparison with other members of the same communities.

Variations in health across racial categories

The legacy of institutionalized discrimination remains a dominant feature of South African society, and many contemporary health statistics continue to be reported according to race or “population group” (2). Despite the ubiquity of this nomenclature, different meanings are attached to racial categories used in epidemiologic and clinical research (102). A common interpretation is the use of race as a proxy for socioeconomic status, with health differences across racial groups continuing to be of interest for monitoring inequities and facilitating the appropriate distribution of public resources (103). In other situations, race may also be interpreted as representing sociocultural practices or health beliefs (104). In addition, racial differences in health outcomes are sometimes attributed to genetic differences (105), with care being taken not to echo the early 20th century attribution of disease differentials to innate qualities.

In a critique of these current practices, some writers have argued that the continued use of these classifications in South Africa serves to reinforce the concept of “race” as a biologic or cultural category and have called for an end to the disaggregation of health statistics by racial group (106–108). In response, other researchers have defended the continued use of racially disaggregated statistics (105, 109). This debate is reflected in variability in the use of racial classifications in public health circles in South Africa. For instance, racial categories have been removed from death certificates, but they remain on official census forms and are a common feature in government-supported public health research.

Disentangling the effects of socioeconomic position from racial categories is an important challenge in social epidemiologic research in South Africa, as in the United States and elsewhere (110). To date, few studies have investigated these issues (50). However, the availability for the first time of comprehensive national data sets (the first nationwide census seeking to enumerate all South Africans took place in 1996, and the first Demographic and Health Survey was conducted in 1998), as well as long-term cohort studies of health and development (111, 112), presents significant opportunities. Questions that may now be examined include how different measures of socioeconomic status are intercorrelated within population groups and whether the associations between socioeconomic status and various health outcomes vary across population groups.

Occupation and health

Occupation continues to be a central topic in social epidemiology in South Africa. Mining has received ongoing attention as the most hazardous industry in the country, with
sustained high rates of occupational trauma, silicosis, and tuberculosis. The migrant labor system, which originated in the mining industry and was reinforced by apartheid-era laws, required that workers from rural areas work in multiple short-term jobs and return to their homes when they were between contracts, ill, or no longer able to find work. The effect was both to delay the onset of silicosis and to make occupational illnesses largely invisible among active miners (113). This labor pattern, coupled with neglect by the industry, hindered research into occupational diseases of long latency in this population. As a result, there were almost no studies of silicosis among Black miners until the 1990s (114), when predictions of a growing epidemic of lung disease emerged (115). These predictions were based on data showing that silica dust concentrations in the mines had not changed for a number of decades, as well as the stabilization of the migrant labor system, with workers being employed in the mines under continuous long-term contracts (116). The paradoxical result of this improvement in conditions of employment has been an increase in disease among miners: Because of prolonged exposure to occupational hazards, the incidence of silicosis and related diseases appears to be on the rise.

Given the long latency period of lung disease, the effects of changing employment practices in the mines are only now becoming evident in the increased burden of lung disease among older active and retired miners. Recent research among ex-miners in the labor-sending areas of Botswana and the Transkei found occupational lung disease in approximately 25 percent of the men (117, 118), which suggests a huge “externalized” burden of disease. Despite attacks on the quality of these studies by the mining industry (119), the existence of an ongoing epidemic of silicosis in the gold mines has been confirmed among older in-service miners (120). In addition, “mining-related lung disease” has a new meaning in South African epidemiology at the turn of the 21st century: The potent combination of silicosis and HIV infection has been shown to dramatically increase the risk of tuberculosis among miners (121).

The investigation of occupation as a socioeconomic determinant of health in South Africa must also account for the impact of unemployment. As in numerous other low- and middle-income nations, levels of unemployment across the country are over 30 percent and are likely to have a strong influence on population health (48). Epidemiologists are beginning to investigate the health impact of the rapid growth of marginal employment and the informal job sector (122–126). Finally, the effects of global economic changes on the health of South Africans may be substantial, and the task of measuring and understanding the impact of these changes is a significant challenge both in South Africa and elsewhere.

THE HIV/AIDS EPIDEMIC

HIV/AIDS is the principal health threat to South Africa, now and for the foreseeable future. According to antenatal clinic surveys, the national seroprevalence of HIV has increased from approximately 1 percent in 1990 to more than 20 percent in 2000 (63). Recent modeling of the epidemic suggests that by the year 2010, one sixth of the country’s population, or about eight million people, will be infected with HIV (127). On the basis of the present trends, the Joint United Nations Programme on HIV/AIDS estimates that a 15-year-old boy living today in South Africa has approximately a 60 percent probability of becoming infected with HIV and dying of AIDS (128).

Superimposed on the political and economic changes the country has experienced in the past decade, HIV/AIDS is beginning to have devastating effects on South African society (129). It is expected that by the year 2010, more than 15 percent of all South African children under the age of 15 years will be orphaned, the vast majority of them because of AIDS (130), and that death rates among economically active adults will almost double, changing the demographic profile of the country and dramatically slowing population growth (63). In this context, social epidemiologists in South Africa have an essential role to play in bringing the epidemic under control and addressing its consequences. Below we draw attention to some of the roles social epidemiologists might play in understanding and responding to this public health catastrophe (for further discussion, see Myer et al. (17)).

Migration and HIV/AIDS

As in other parts of the world, there is widespread evidence for a role of labor migration and other forms of population mobility in the dissemination of HIV across South Africa (131, 132). Typically, male migrant workers have multiple sexual partners while employed in urban areas and then return to wives and families at their rural homes (133, 134). These migrants provide a bridge for the spread of HIV infection from high-risk female sex workers to women in rural communities, who appear to have a low risk profile in traditional terms (135, 136). Although the general association has been described repeatedly in different South African populations, there are still few data on how different forms of migration may affect the dissemination of disease at the population level and, more importantly, what steps may be taken to reverse the apparent pattern. In addressing these questions, epidemiologists are beginning to turn to research on the structure of social and sexual networks within and between populations (137–139). The combination of empirical social research with epidemiologic modeling of the transmission of infectious diseases holds promise for future understanding of the population-level spread of HIV. In the case of South Africa, special questions emerge regarding the factors that shaped the geographic and temporal dissemination of the epidemic during the 1990s, particularly how the social and economic processes of urbanization affected the spread of the disease.

Social change and HIV/AIDS

In South Africa and elsewhere, epidemiologic analyses of the HIV/AIDS epidemic have given insufficient attention to the role of social change and its potential importance in prevention efforts (140, 141). While some analysts have commented that the magnitude and impact of the epidemic may be mediated in part by the general social and economic factors that shaped the geographic and temporal dissemination of the epidemic during the 1990s, particularly how the social and economic processes of urbanization affected the spread of the disease.
patterns of South African society (142), there has been little investigation into the specific social processes that may contribute to this. For instance, questions emerge about how the spread of infection is affected by popular stigmatization of HIV disease, since the widespread fear of discussing HIV infection status may inhibit the effect of both primary and secondary prevention efforts (143, 144). Meanwhile, social capital and other shared resources within communities may serve to both buttress prevention efforts and moderate the impact of the disease on affected households (145, 146). In addition, programs making antiretroviral therapy for HIV/AIDS available may have an important impact on the social norms associated with HIV disease. The shift of HIV/AIDS from a painful and universally fatal condition to a chronic disease that can be managed over the long term may have positive impacts on the prevalence of stigma and on social support systems (147, 148). The fact that these issues have not received adequate attention thus far is perhaps unsurprising, given the disciplinary divides that social epidemiologists constantly seek to cross (149). While investigation of the determinants of HIV infection falls within the traditional purview of epidemiologists, topics such as stigmatization and social movements remain within the domain of social scientists (150). Social epidemiologists will need to bridge these gaps if we are to understand how specific social phenomena affect the course of the HIV/AIDS epidemic.

Social epidemiology and the rhetoric of AIDS denial

The public discourse on HIV/AIDS has a unique history in South Africa that touches on social epidemiology in particular. President Thabo Mbeki’s assertion that HIV is not the cause of AIDS has been widely publicized (151–154) and has been the subject of much speculation (155–157). Social and economic determinants have been prominent features of Mbeki’s alternative explanations (158), and he frequently describes AIDS in South Africa as a disease of “poverty and underdevelopment” (159). Mbeki’s Minister of Health, Dr. Manto Tshabala-Msimang, follows this position and commonly speaks on the role of socioeconomic deprivation as a cause of HIV/AIDS: “The determinants of the epidemic cannot be explained only in terms of individual risk-taking behavior, but are varied and complex. Some of the causal factors are poverty, social deprivation, and the economic alienation suffered by most South Africans” (160). Taken out of context, such quotes would suggest a relatively sophisticated understanding of the socioeconomic determinants of HIV/AIDS. However, using these arguments, the Minister has promoted indigenous plants and nutritional supplements in place of antiretroviral therapy for the treatment of HIV infection (161), a position that gave way only in 2003 in response to overwhelming domestic and international political pressure.

Although the rhetoric of AIDS denial over the past 5 years in South Africa has drawn heavily from positions that are central to social epidemiology, this perspective misappropriates the view of many social epidemiologists, who would accept that the causal pathways involved are multileveled and thus incorporate both socioeconomic determinants and viral etiology. However, paradoxically, the backlash against Mbeki’s and Tshabala-Msimang’s positions may have led to an overshadowing in public health circles of societal factors as determinants of the spread of HIV/AIDS, leading many domestic researchers to focus on more proximal behavioral and biologic determinants of the epidemic.

Significantly, the most vocal criticism of Mbeki’s views on HIV/AIDS has come from a domestic nongovernmental organization, the Treatment Action Campaign, which has taken successful legal action and led various popular protests against the national government’s position. In drawing widespread support and providing an effective voice for the millions of HIV-infected South Africans (157, 162), the Treatment Action Campaign has provided an important local example of how social mobilization can contribute to improvement in the health of populations. The role of the Treatment Action Campaign in changing norms and perceptions regarding HIV-infected persons, and in turn possibly affecting the course of the HIV/AIDS epidemic in South Africa, requires greater consideration from epidemiologists and social scientists alike.

CONCLUSION

The evolution of social epidemiology in South Africa offers a glimpse of the complexities that arise as the discipline extends to populations around the globe. Evidence from South Africa suggests that the links between socioeconomic factors and health are perhaps more dynamic and varied than may be commonly assumed in the developed world. Given this complexity, epidemiologists are challenged to investigate how social, economic, and political processes (both historical and contemporary) combine to shape the different social and economic features of a society and how these features, in turn, affect health. The persistence of unacceptable disparities in socioeconomic status and health despite the transition to a democratic postapartheid polity means that the insights of social epidemiologists are likely to continue to play an important role in South African public health.

The example of South Africa also provides a useful vantage point from which to consider some current debates in epidemiology. While many observers have noted that epidemiologic research is not entirely apolitical (163, 164), the degree to which epidemiologists should be engaged in public health debates and broader social changes remains controversial. The history of social epidemiology in South Africa indicates that it is possible to contribute to sociopolitical change by documenting the nature and causes of health inequalities. The example of South Africa also suggests that in some instances, when epidemiologists overlook social and economic contexts, they may run the risk of appearing to be complicit in social inequities. In some respects, it may be argued that this dilemma remains pertinent today in the context of the dramatic inequalities in global health.

Finally, it is widely recognized that the HIV/AIDS pandemic presents an epic challenge to epidemiologists and that the South African experience will be important in gauging the success or failure of the international public health response. However, the significant role that social epidemiologists can play in confronting HIV/AIDS is not
always appreciated. Although its full social and economic impact is only beginning to emerge, it is already clear that the HIV/AIDS epidemic will be devastating in many areas of South African life. An effective response will require integration of epidemiologic research on both socioeconomic factors and infectious disease processes—without an exclusive focus on either.

ACKNOWLEDGMENTS

The authors thank Chelsea Morroni and Dr. Alan Berkman for their comments on various drafts of the manuscript and Dr. Di Macintyre for her assistance in identifying sources of unpublished information.

REFERENCES


