CONFERENCIA MAGISTRAL

THE SOCIAL ROOTS OF HUMAN DISEASE: THE ORIGINS OF HEALTH DISPARITIES BY RACE AND CLASS*

Leon Eisenberg**

The distribution of health and disease in human populations reflects where people live, when in history they live, the air they breathe and the water they drink, what and how much they eat, the energy they expend, the occupations they have, the status they occupy in the social order of their communities, whether they are socially isolated or surrounded by friends, and the quality and the amount of medical care they receive. None of this will be news to the audience, except perhaps the order in which the relevant variables are listed, with medical care listed last. Can that be correct? Think about that question in the following context.

Mortality from respiratory tuberculosis, as high as 400 per million in England and Wales in 1840, fell by half by 1880, two years before Koch identified the TB bacillus. By 1940, before there was any effective medical treatment, tuberculosis mortality had fallen to 10% of its level a century earlier (McKeown, 1975). The healthcare system had little to do with these profound changes; they reflected better housing, improved nutrition, less hazardous working conditions and public education. Only in the last 50 years has medical care become increasingly decisive in clinical outcomes when care is responsive to community needs.

The thesis I propose to develop is that social variables are decisive in determining, not only what diseases strike which persons, but also in modifying the very biology of the disease process. To do so, I will make five points:

ONE: Health disparities parallel economic and social disparities.
TWO: It is not solely the national product per capita, but the distribution of income within populations that matters for health.
THREE: Medical technology, to the degree that it is effective, will increase rather than decrease health disparities between social classes because of differential access.
FOUR: Changing social conditions, including patient behaviors, improvements in medical technology, and the provision of medical care alter the very biology of disease itself.
FIVE: Finally, I will argue that, if health disparities are to be reduced, physicians must move beyond their clinical roles to become social activists.

Health disparities parallel social disparities

The great pathologist Rudolf Virchow (Eisenberg, 1984) proclaimed, a century and a half ago, that the measures taken to combat disease must be as much social as they are medical because the social environment is decisive for health. Physicians, he said, are the natural attorneys of the poor. Progress was to be measured by vital statistics. In his words:

“We will weigh life for life and see where the dead lie thicker, among the workers or among the privileged.”

Where do the dead lie thicker? Epidemiologic data

**Maude & Lillian Presley Professor of Social Medicine and Professor of Psychiatry, Emeritus. Harvard Medical School, 641 Huntington Avenue, Boston, MA 02115-6019. E-Mail: Leon_Eisenberg@HMS.Harvard.edu
are unequivocal: health disparities parallel economic and social disparities. The World Health Organization has created a comprehensive statistic, Disability Adjusted Life Years (DALYs), to measure disease burden more fully. The DALY takes into account not only differential mortality rates, but the age at which each death occurs and the disability produced by disease (that is, it computes years of life lost by premature death and discounts the value of years lived with disability). How are DALYs distributed among nations? DALYs lost per thousand population are five times greater in sub-Saharan Africa than they are in Scandinavia, Japan, or the U.S.

Consider a second data point: deaths under age five. In 2003, 10 million children will die of preventable and treatable diseases. Mortality in under-fives will be highest in sub-Saharan Africa. Malnutrition, diarrhea, measles, pneumonia, malaria, and birth asphyxia are the major killers. The children will die because their communities lack clean water, sewage disposal and they have not been immunized; many are orphaned because their mothers lack access to obstetrical care; their protein and micronutrient intake are insufficient and they lack access to insecticide-impregnated bed netting, antibiotics for infections, and oral rehydration therapy (Black et al, 2003; Bryce et al, 2003; Jones et al, 2003; Bellagio Study Group on Child Survival, 2003).

Consider a third datum. Mental disorders exact an enormous health burden (World Health Report, 2001). That burden is differentially distributed. Since Edward Jarvis (1855) surveyed “insanity and idiocy” in Massachusetts in the mid-19th century it has been known that, to use his quaint English, “the pauper class furnishes [many more] ... cases of insanity than the independent class”. Despite changes in the definitions of mental disorders and in the methods of study, epidemiologic research verifies his finding. In a review of the problem in developing countries, Patel and Kleinman (2003) note a consistent association between poverty and mental ill health, which they attribute to the insecurity and hopelessness, the risks of violence, and the high rates of physical ill health that accompany living in a chaotic social environment. Primary prevention is possible, at least in theory, by improving employment opportunities, housing, and healthcare in impoverished environments. Surprisingly enough, treatment has been shown to be effective. This has been shown for individual and group psychotherapy among low-income women with depressive symptoms in Mexico City by Lara and her colleagues (2003) at the National Institute of Psychiatry Ramon de la Fuente, by Miranda et al (2003) in Los Angeles through either medication or psychotherapy, and by Bolton et al (2003) for group interpersonal psychotherapy for depressed patients in rural Uganda. What we have encountered is preventable and treatable mental illness.

The distribution of income matters

Health status increases as income and education increase, whatever index of health is used (life expectancy, mortality, morbidity, disability, or self-perceived health status). The slope of the curve is steepest under conditions of severe deprivation. This finding holds true when countries are compared with each other and when comparisons are made within single countries. It reflects a fundamental underlying factor: the distribution of incomes within societies influences the average level of health (Wilkinson, 1996). That is, the greater the disparity between the well off and the poor, the lower the mean life expectancy at a given GNP. This was manifest in cross-sectional data from 56 rich and poor nations (Rodgers, 1979) and in comparisons among the 11 relatively well-off countries belonging to the Organization for Economic Cooperation and Development (Wilkinson, 1992).

Per capita conceals as much as it reveals. Brazil is classified by the World Bank as an upper middle income country; yet 1 in 6 Brazilians earns less than $1.00 a day (World Bank). The top 1% of US families own 38% of the nation’s household wealth; they have seen their wealth increase by 42% during the past 20 years; during the same interval the poorest 40% of families lost 76% of theirs (Wolff, 2002). It will not surprise you to learn that the number of Americans living below the poverty line rose by more than 1.3 million last year to a total of 34.8 million (Clemetson, 2003). What are the consequences of this enormously skewed distribution?

Lynch and his colleagues (1998) examined income inequality and mortality in 282 metropolitan areas within the United States. Death rates in metropolitan areas that had the greatest income inequalities were far higher than in areas with less spread in income between upper and lower quintiles. The amount of excess mortality was not a minor statistical finding; it ranged from 65 to 96 per 100,000. Projecting those mortality differences onto the US as a whole, Lynch found that they are equal to the combined loss of life from lung cancer, diabetes, motor vehicle crashes, HIV infection, suicide, and homicide!

In non-industrialized countries, health is not simply a matter of income. When the nations of the world are ranked according to infant mortality rates (IMR), the ranking is inversely correlated with GNP per capita (the lower the GNP, the higher the IMR). But there
are notable exceptions; that is, countries with a lower IMR than expected from their GNP and others with higher mortality. Superior achievers include Costa Rica, Jamaica, and Sri Lanka; poorer achievers Saudi Arabia, Oman, and Iraq. The average IMR for the first three countries is half that for the second three; even though the Arab countries have 10 times the GNP per capita.

What accounts for the remarkable discrepancy? In the countries where infant mortality is low, 90% of the eligible young girls are enrolled in elementary school; enrollment is less than 50% in the higher mortality countries. The correlation coefficient between percentage of female primary school attenders and IMR (0.8) is far greater than that between GNP and IMR (0.3) (Caldwell, 1986). School attendance is a proxy for an interconnected set of social values. A society that educates girls as well as boys is a society which acknowledge women's rights, provides better healthcare for mothers and children, and is sensitive to reproductive rights and family life. Educated women marry later, marry better-educated men, have longer interbirth intervals, and provide better care for their children. The status of women is a sensitive index of child health. Where women cannot breathe easily, children are sure to suffocate.

Medical technology magnifies health disparities

Can health disparities be reduced by the powerful new technologies medicine now possesses? The fact is that technologies increase disparities rather than decrease them, precisely because they are effective, even when they benefit both poor and the rich (Wise 2003). Why? Should that be so? It results from differential access. Please understand that this is not an argument against technology; it is an argument for ensuring universal access! I illustrate that point with five examples.

• Neonatal mortality rates have declined substantially for both black and white infants in the US over the past 20 years because neonatal intensive care units are effective. Yet, the ratio of Black to White neonatal deaths has increased over that period of time. Why? Because of problems of access among the poor (Wise, 2003).

• With modern treatment methods, mortality from acute lymphoblastic leukemia has been reduced from 90% to 20% in the US during the past two decades (Silverman, 2003). Yet it is still 70% in Honduras because of barriers to treatment and limited availability of comprehensive care (Lilleyman, 2003; Metzger et al, 2003).

• Highly active antiviral treatment and antibiotics for secondary infection have revolutionized the epidemiology and clinical course of HIV/AIDS. That statement applies to patients in the US and Western Europe. HIV positive patients in South Africa suffer the ravages of AIDS as it was seen in the West three decades ago.

• A final example illustrates the complexities associated with access to technology. Cesarean sections reduce morbidity and mortality in cases of cephalo-pelvic disproportion, transverse position, placenta previa, toxemia of pregnancy, and fetal hypoxia. These complications are much more common among poor than among middle-class women. Yet, in 1997, the rate of C-sections performed in private hospitals in Mexico was twice that in public hospitals. The same is true in Brazil, where C-section rates exceed 70% in private hospitals and maternity clinics (Finger, 2003). According to the World Health Organization, C-sections undertaken for the convenience of the pregnant woman or the obstetrician put mother and fetus at greater risk, a point some dispute (Minkoff and Chervenak, 2003). There is a virtual iatrogenic “epidemic” of C-sections among the wealthy while the poor who stand to benefit from surgery have limited access to it (Castro et al, 1993).

What Virchow said about disease epidemics in the 19th century can be said about the 21st.

"If disease is an expression of individual life under unfavorable conditions, then epidemics must be indicative of mass disturbances of mass life."

The social construction of clinical diabetes

The evolution of diabetes as a clinical entity graphically illustrates the interactions among heredity, mode of life, and means of care, on the one hand, and the biology of disease process itself, on the other. A once acute and fatal disease has been transformed into a chronic, debilitating disorder, one that targets the rich in poor countries and the poor in rich countries; diabetes is particularly savage to the chronically poor who become suddenly relatively rich.

Diabetes mellitus is characterized by defective regulation of glucose metabolism. There are two principal forms: insulin-dependent (Type I) and non-insulin dependent (Type II). The first is relatively uncommon, the latter some 15 times more common. Persons with Type I have an absolute deficiency of insulin secretion associated with pancreatic islet atrophy; Type II patients suffer from tissue resistance to relatively normal amounts of secreted insulin.

Although Type I diabetes has been found among all
populations studied, its incidence varies almost 60-fold between countries. For example, within Italy alone, rates are 30 per 100,000 in Sardinia (the second highest incidence in the world) versus 6.5 per 100,000 in the Lazio region of the Italian mainland, a region lying opposite Sardinia across the Tyrrhenian Sea. Sardinians are a relatively homogeneous population, genetically distinct from other Italians. Historically, there had been little exchange between Sardinia and Lazio until 1950 when many Sardinians began to settle in Lazio as the result of post-war economic opportunity. Mutoni and colleagues (1997) compared the incidence of IDDM in children born in Lazio to parents of Sardinian origin. The incidence of IDDM among children born in Lazio of two Sardinian parents is four times as high, and among children of mixed marriages twice as high, as the rate among indigenous children, pointing to a strong genetic component.

But there is more to the story than this. There has been a steady increase in the incidence of IDDM in Sardinia itself over the past several decades, indicating an as yet unidentified environmental agent. The increase in prevalence reflects the remarkable success in treating diabetes; higher prevalence, reflecting greater survival of child and adolescent cases, leads to further increase in incidence as more Type I diabetics survive to produce viable offspring.

Type II diabetes also has an hereditary component as evident from (a) greater concordance in identical twins, (b) aggregation in families and (c) marked differences between geographically and ethnically separate populations. Children of parents with NIDDM, later to become diabetic, exhibit hyperinsulinemia on oral glucose tolerance testing a decade or more before hyperglycemia appears. At the same time, an environmental contribution is evident from higher risk for clinical disease with (a) lower activity level, (b) higher caloric intake, and (c) more obesity.

The prevalence of diabetes is increasing, both because diabetics live longer and because incidence is rising. Race and class are determinants of disease outcome. Death rates for diabetes in the US are 2.5 times greater among African-Americans and American Indians and 1.7 times greater among Hispanic Americans than they are among Whites. Mortality from diabetes is inversely correlated with income. Women in families earning less than $10,000 have three times the death rate of women in families earning more than $25,000; the ratio for men is 2.6 (Health, US, 1998, p. 96). Diabetes leads to a three times greater loss of years of potential life among African-Americans than among Whites in the US, a measure of racism in US society.

Incidence rates for diabetes have been increasing in parallel with obesity and physical inactivity in the UK (Bagust et al, 2002), Australia (Dunstan et al, 2002), and the US (Mokdad, 2003; Ogden et al, 2003). The time interval is far too short for the genes contributing to risk to have increased in frequency. What has changed is the mode of life.

Of particular interest are the "epidemics" of diabetes that have appeared among Polynesians, American Indians, and Aboriginal Australians as their lifestyles have been «modernized.» A striking recent example occurred on Nauru, a small Pacific Island inhabited by about 5,000 Micronesians. Until World War II, high-energy expenditure was required for sheer survival via fishing and subsistence farming. After the war, phosphate mining by foreign companies yielded rental income for the Nauruans that rapidly transformed them into one of the world’s wealthiest and most sedentary peoples. Today, virtually all foodstuffs are imported; most have a high caloric content; obesity is ubiquitous. NIDDM, previously minimal, began to reach epidemic proportions in the 1950s and afflicted almost two-thirds of 55 to 64 year old adults. Paradoxically, Nauruans acquired one of the world’s shortest life spans from diabetes and its complications because of rising income and its sociobehavioral consequence.

The distribution of the disease among Nauruans has continued to change during the past 50 years. Health surveys in recent decades reveal that the age standardized prevalence of impaired glucose tolerance peaked at 21% in the mid-70s and then declined to half that value by the late ’80s. The most parsimonious explanation for the rise and subsequent fall is that diabetes resulting from the affluent life-style has already afflicted the genetically most susceptible Nauruans, leaving a residual population of relatively resistant individuals. The Nauru epidemic has ominous implications for Southeast Asia. Rates of diabetes among Chinese and Indian expatriates living in the West (in contrast to low rates in China and India) makes it likely that the improved living standards anticipated for India and China in the next century will lead to epidemics of Type II diabetes.

How did the NIDDM genotype become widespread? Higher mortality and shorter longevity should have led to adverse genetic selection. Professor James Neel has proposed the "thrifty genotype" hypothesis. During most of our history as a species, human life has been characterized by fluctuating food supply and frequent famines. A quick insulin trigger reduces calorie loss and permits more fat storage during periods of relative plenty; insulin resistance in muscle may also contribute to the «thrifty genotype» by blunting the hypoglycemia associated with fasting. Individuals with thrifty adaptations (i.e., those able to
release insulin rapidly when a temporary food glut becomes available) can convert most of their ingested calories into fat. Greater fat stores would make them better able to survive subsequent periods of starvation.

The very same genotype becomes a handicap in the presence of abundant high calorie foodstuffs and reduced physical activity.

Until World War II, the population on Nauru was under intense pressure for selection of the thrifty genotype; their ancestors had reached the island only after long sea voyages; crop failures on the island were common (indeed, many Nauruans suffered from starvation during the Japanese occupation). The sudden change in economic circumstances on Nauru created the conditions for an "epidemic" The creasing and recession of the epidemic display in heightened fashion was what occurred over a century in the West on a more gradual course with a longer period of accommodation. Neel has pointed out is the biological entity has been converted from an acute and uniformly fatal affliction into a chronic ailment whose secondary complications have come to dominate the patient's life experience. Elliott Joslin, one of the first Boston specialists to treat diabetic patients, summarized progress during the pre-insulin era in his 1922 Shattuck lecture: "The average known duration of the fatal cases of diabetes in the city of Boston between 1895 and 1913 was 3.3 years; during 1915, it was 4.3 years and 1920, it was 5.3 years."

In the early years of clinical management, diabetics died of the acute complications of their disease: coma, gangrene, and infections. Joslin's scrupulous attention to hydration, to diet, and to personal hygiene partially controlled these complications. Nonetheless, in the pre-insulin era, half of IDDM patients died within 20 months of the diagnosis; less than 1 in 10 survived as long as five years. Joslin hailed the discovery of insulin as "the advent of the promised land" for diabetics. His triumphalism proved to be premature.

Thirty years after the introduction of insulin, Joslin reported that the average age at death had risen from 44 to 64 years. Diabetic coma, as a cause of death, had fallen from 64% to less than 2 percent, but cardiovascular and renal deaths had risen from 17 to 70 percent. In the last 50 years, the addition of antibiotics, antihypertensive treatment, renal dialysis and vascular surgery to a comprehensive program of care has prolonged survival still further, but at the cost of retinopathy, nephropathy and vascular complications (coronary heart disease, stroke, and peripheral vascular disease). I do not want to be mistaken. The gains are unmistakable. Few diabetics would choose to return to the pre-insulin era. I stress the unforeseen biological consequences of medical progress in order to highlight the importance of continuing research into the pathophysiology of diabetes.

Progress in clinical management continues. For many years, a bitter battle was fought between protagonists for precise and close control of blood sugar levels and others who championed a «liberal» regime. Decisive evidence for the superiority of one or the other mode of management was not available until the completion of the Diabetes Control and Complications Trial. It enrolled 1400 patients with Type I diabetes in a multicenter study comparing standard with intensive care. Patients in the intensive care arm of the study were placed on one of two regimens: multiple daily insulin injections or continuous subcutaneous insulin infusion by pump. Treatment was initiated by four days in hospital, followed by frequent individual outpatient visits, group meetings, and telephone calls to review progress in order to monitor hemoglobin A1c levels, adjust insulin dose, maintain diet and weight control, and regulate exercise patterns.

The results were unequivocal: the intensive management program significantly delayed the onset and slowed the progression of the microvascular and neurologic complications of diabetes; the one important side effect was a modest increase in the number of hypoglycemic episodes. The annual cost of intensive treatment was 3 times higher than that for conventional care. For that additional cost, each patient, on average, gained an additional 7.7 years of sight, 5.8 years free from end-stage renal disease and 5.6 years free from lower extremity amputation.

The powerful benefit of close and continuous involvement with patients and active patient participation in managing chronic disease is evident from a datum the authors present as an incidental finding: ninety-nine percent of the patients completed the trial! This extraordinary rate of "compliance" with care is a tribute to the cooperative relationship between the research team and its patients; the treatment program demanded major lifestyle changes and strict adherence to a demanding protocol. Patients had to understand what they were doing and why; nurse clinicians were available for consultation when needed in addition to the regular phone and clinic visits they provided. Clinicians and patients were enrolled in a common enterprise; patient education was part of the protocol; and availability of unhurried consultations was assured. In that clinical context,
intensive care worked; without the psychosocial component, adherence would assuredly have been far lower.

Physicians as social activists

To this point, I have presented evidence that health disparities correlate with socioeconomic disparities and that advances in medical technology, useful as they may be for rich and poor alike, cannot, by their very nature, eliminate class-based health inequities. I conclude that physicians have a moral imperative to serve as social advocates for the poor and the dispossessed because of what we know. We in better-off countries must help to persuade our governments and fellow citizens to invest a greater proportion of our resources in improving life chances for the poor. It should be done, in the first place, because it is the right thing to do. In this case, what is right is also in our self-interest. Infections do not respect national borders. Multi-drug-resistant tuberculosis bacilli, human immunodeficiency virus, and SARS do not await visas to cross borders. Investment in human development is essential to reduce the poverty and despair that breed war and terrorism.

In each country, physicians must cultivate allies in the struggle to reduce health inequities. Policies and the politics will differ from country to country. I would not presume to offer prescriptions for Mexico. In my country, it is unconscionable that one in six Americans lack health insurance and have no regular source of care, despite the highest healthcare expenditures per capita are in the world. Recently some 8,000 US physicians signed a manifesto calling for legislation to capita are in the world. Recently some 8,000 US physicians signed a manifesto calling for legislation to create tax-supported, universal health insurance. Physicians as social activists...