

Theories of disease causation: Social epidemiology and epidemiological transition

Soma Hewa

School of Policy Studies, Queen's University, Kingston, Ontario, Canada and School of Public Administration, Guangzhou University, Guangzhou, China.

*Correspondence: Professor Soma Hewa
E-mail: hewas@queensu.ca*

Abstract

Modern Western medicine is based on scientific reductionism, the key assumption of which is that complex phenomenon can be understood by reducing and dividing into smaller parts. The “biomedical model” diagnoses disease in terms of measurable biological parameters, and treats the patient as a biological organism. Although this particular approach has been responsible for tremendous successes in modern medicine, the tendency of the biomedical model to neglect social, economic and psychological factors has been criticized, and alternative explanations have been sought to complement it. This paper presents a brief analysis of these models.

Key words: *Social epidemiology, germ theory, multi-causality, general susceptibility*

Introduction

The foundation of modern Western medicine is generally described as the biomedical model. It assumes that diseases are the results of deviations in the normal functioning of biological (somatic) variables. Further, it distinguishes between the body and the mind, and asserts that disease can be treated separately from the mind. As diseases represent some form of departure from the normal order of biological variables, they can be cured by medicine, which restores the normal functioning of the body. The body is thus like a machine, and any breakdown of it a disease can be repaired. Since medicine takes the mechanical metaphor, it presumes that the doctor is a mechanic who undertakes the task of repairing the dysfunctioning body. As the whole biomedical model the body, disease, medicine, and doctor has been understood in terms of a mechanical metaphor, it is described as the “mechanistic biomedical model” (1).

This particular approach to the human body tends to overestimate the role of medicine as a technological imperative. Consequently, it leaves no room within

its framework for social, environmental, behavioural, and psychological dimensions of health and illness. The biomedical model not only requires that the disease be dealt with as an entity independent of social and behavioural factors, it also advocates that behavioural disorders themselves are caused by bio-chemical or neuropsychological deficiencies. Subjective interpretations of health and illness are deemed irrelevant.

Clearly, biomedicine takes a reductionist approach to disease. The complex phenomenon of disease is reduced to a single primary principle of cause and effect. Here, the primary causal factor is physical, which can be explained in terms of the language of chemistry and physics. The philosophical foundation of the mechanistic biomedical model was reinforced by the development of the *germ theory* of disease in the nineteenth century, which postulated that every disease was caused by a specific, identifiable agent (such as bacterium or virus). Cures, which were understood to destroy the infecting agent within the biological realm, came to be sought. In a series of empirical studies between

1861 and 1880, Louis Pasteur demonstrated that the injection of *cholera vibrio* into chickens produced this (and no other) disease. In 1882, Robert Koch discovered the *tubercle bacillus*, which was pronounced as the cause of tuberculosis. These two important discoveries contributed to the doctrine of a “specific aetiology” and the concept of a “magic bullet” (chemical agents that destroy the organisms). Researchers strived to find single, specific causes of disease, and then destroy them with “magic bullets,” leaving other organisms unharmed (2).

Although the movement of social medicine was active in the late nineteenth century, it was seen as too radical because it advocated social reforms for improving the living conditions of people. For example, Rudolf Virchow advocated developing both medical and economic conditions in Germany, and wanted to bring about social reforms affecting living conditions to complement the emerging bacteriological discoveries. Edwin Chadwick, Friedrich Engels, and John Snow exposed the terrible living and working conditions of the people that contributed to widespread diseases during the industrial revolution. Engels, in one of his best known essays, *The Condition of the Working Class in England* (1845), wrote: “When one individual inflicts bodily injury upon another, such injury that death results, we call the deed manslaughter; when the assailant knew in advance that the injury would be fatal, we call his deed murder. But when society places hundreds of proletarians in such a position that they inevitably meet a too early and an unnatural death, one which is quite as much a death by violence as that by the sword or bullet” (3). Such views were often censured as too radical, and were prevented from being absorbed into the mainstream academic discourse and scientific inquiries.

Although germ theory was a major breakthrough in the development of medicine, it inadvertently convinced scientists to focus exclusively on laboratory procedures devoid of social and economic contexts. Instead of making the reasonable judgment that while bacteria and viruses are the prime causative agents of disease, with their pathogenic effects being mediated by socio-environmental conditions, supporters of germ theory totally dismissed the importance of non-biological factors. Thus, the practice of medicine in the twentieth century came to rest solidly on the following

premises: 1) Disease is a process accounted for by deviations from the norm of measurable biological parameters; 2) Disease is best understood by means of a science like pathophysiology, where principles are formulated in terms of molecular biology, biochemistry, and physics; 3) The human patient is a biological organism (“whole body”) such that, in René Descartes' phrase, “were there no mind in it at all, it [the body] would not cease to have the same functions;” 4) Each part of the body is thus a field for specialized knowledge; 5) The cure for disease is typically achieved physically (e.g. reducing, or neutralizing the pathogenic agent); 6) Medicine is an applied science developed through laboratory research (4).

Multi-causality Approach

Although medicine and medical technology continued to be recognized as vitally important contributors to the rising life expectancy and declining mortality in the early twentieth century, the evidence in support of an equally important contribution made by the improved social and economic conditions to health began to gather momentum. Since the 1940s, an increasing number of studies by medical historians and sociologists have shown that only a small part of the decline of mortality from infectious disease in the late nineteenth century was due to bacteriological discoveries, and major diseases such as tuberculosis, dramatically declined long before the development of medicine. For example, Thomas McKeown found convincing evidence that the mortality from infectious diseases, with the exception of the mortality from smallpox, declined since the late nineteenth century with the improvement in nutrition, living conditions, and personal hygiene long before the development of effective medicine. He argued that the infant mortality declined owing to improved nutrition, better care and feeding for infants, and the adequate supply of food for mothers rather than the improved obstetric services. Likewise, the decline of mortality from typhoid and dysentery convinced McKeown that “rapid decline in mortality from diseases spread by water and food since the late nineteenth century owed little to medical intervention” (5).

Notwithstanding the fact that modern medicine and laboratory sciences made significant contribution to the decline of mortality and the increase life expectancy, the process of disease causation is much more complex than a single causal factor to be explained by the germ theory of disease. The decline of mortality with the improved living conditions and public health suggests that disease causation is a complex development, involving a disease agent, a human host and the environment (both social and physical) that need to be understood as a multi-causal process. The disease agents are biological, chemical or physical factors, the presence of which is necessary to cause a disease. The human host, whose personal behaviors, genetic predisposition, immune system, and the emotional state make the person susceptible to the effect of the disease agent. The environment, which is external to the host, contributes either to weaken or to strengthen the host by the societal activities. For example, political upheaval or natural disasters are two different external factors, which have the potential to weaken the “defenses” and to “expose” the human host to disease-causing agents. The individual's *susceptibility* is dependent upon the biological, immunological and psychological capacity to face the hostile environment, and to ward off an attack from a disease-causing agent. The decline of mortality from infectious and parasitic diseases owing to social and economic developments therefore came from at least two different fronts: First, the improvement of sanitary and living conditions, the supply of clean water and sewerage systems significantly reduced the *hostility* of the environment making it “livable” to the human host. Second, the *availability* of adequate quantities of nutritious food, and the access to education improved the knowledge on personal hygiene and feeding practices for infants that strengthened the host's *defensive* capability. Both these developments were clearly outside scientific medicine in the late nineteenth century.

General Susceptibility

The multi-causality approach to disease, like the mono-causal germ theory, highlights the connection between the disease-causing agent and the human host. However, unlike the germ theory, the multi-causality model emphasized the mediating social and physical environment, and the psychosomatic

factors that make human host vulnerable to attacks by disease-causing agents. In this sense, the multi-causal approach seems a step forward in understanding disease and health in their broader socio-environmental context. Yet, the problem with both these approaches is that they could not give an explanation to the relationship between certain demographic variables and mortality patterns in the same population, exposed to the same pathological agents. For example, mortality rates were found to vary between married and unmarried people, social classes, or status groups, and the people with greater social ties or extensive social networks of relationships. As early as 1937, writing on the decline of mortality from tuberculosis in the United States, Wade Hampton Frost argued that the decline of the disease was not confined to the areas where specific measures had been adopted, but far and wide in places where no such measures were taken: “One of the most important factors in the decline of tuberculosis has been progressively increasing human resistance, due to the influence of selective mortality and to environmental improvements such as better nutrition, relief from physical stress, tending to raise what may be called nonspecific resistance” (6).

Taking the queue from this early investigation into the hypothesis of improved *nonspecific* host resistance, social epidemiologists, Lisa Berkman and Cassel Syme presented a 9-year perspective study of 6,928 adults living in Alameda County in California. The result showed that the people with fewer social ties at the time of the initial study in 1965 were more likely to die over the 9-year follow-up period. An index of social ties (that measure the degree of social integration), which included a range of social connections such as marital status, contacts with friends and relatives, organizational membership, and church membership, was correlated with mortality rates from the first study in 1965 to the 9-year follow-up period. What they found was astounding. Those who were most isolated were 2 to 3 times more likely to die from a wide range of causes, such as heart diseases, cancer, respiratory disease, gastrointestinal, and all the other causes of death during the follow-up period, compared to those who were the most socially connected. The association between social ties and mortality was found to be independent of

self-reported physical health status during the initial study in 1965, socio-economic status, and health practices, such as smoking, alcohol consumption, obesity, physical activity, and utilization of preventive health services (7).

Over the years, since the publication of this study, several others who carried out similar research concluded that people who are more socially connected: 1) live longer; 2) are more likely to survive a myocardial infarction; 3) are less likely to experience a recurrence of cancer; 4) are less likely to suffer from infectious illness, than those who are less integrated to the community (8). What these findings highlight should not be entirely new to those who are familiar with the work of the classical sociologist, Emile Durkheim, who studied suicide in Europe in the late nineteenth century. He clearly established the link between suicide rates and the level of social integration. Although suicide is clearly a *personal* act of self-destruction, it is deeply influenced by certain social conditions that the individual has come to reject. For Durkheim suicide was not merely a personal tragedy, but a condition of an entire society that has a tragic personal consequence (9).

Social epidemiologists have established that social isolation often manifests as certain behavioral modifications of the individual that leads to unhealthy lifestyle choices, such as smoking, drinking, overindulging and the lack of physical exercise. Robert Putnam who carried out several studies on social capital argues that: "As a rule of thumb, if you belong to no groups but decided to join one, you cut your risk of dying over the next year *in half*. If you smoke and belong to no groups, it's a toss-up statistically whether you should stop smoking or start joining" (10). He argued that social connectedness is the most basic, but the most powerful determinant of human wellbeing. Perhaps this is what classical political philosophers (like Marx and Hegel) meant when they wrote: "human beings are social beings." To negate this very human nature comes at a price.

In essence, the question that social epidemiologists were trying to answer was: "how do social relations get inside of our bodies." People with extensive social networks receive material and emotional support during the time of highly stressful situations in their lives. Such community supports and the sense of belongingness function as a "safety net" that

prevents the individual from sinking into despair. This reduces the *general susceptibility* to disease, rather than to any specific disorder. In addition, social relations also have positive psychological impact on the individual's behavior through enhanced motivation to be more socially active. Social epidemiologists have studied the direct biological link between social support and enhanced host resistance. What they discovered was the clear connection between neuroimmunological endocrines and stress. Most everyday stressful experiences activate multiple hormones, and affect the functioning of multiple biological systems, and could cause serious end organ damage (heart, kidney, brain etc.) by the prolongation of the situation. Research has shown that protracted stress leads to rapid aging mediated by neuroimmunological endocrines. Stress affects health multiple ways mediated by both external and internal reactions: while *internal* mediation could activate neuroimmunological endocrines, *external* reactions to stress may lead to behavioral modifications, such as drinking, smoking, overeating and the lack of physical exercise. Both these processes mediate between the persistent stressful experiences and the biological manifestation of illnesses such as heart disease, cancer, stroke, peptic ulcer, and infectious diseases (11).

The general susceptibility thesis explains the level of *vulnerability* of a given population (social group) to diseases. The *Whitehall* study reinforced the general susceptibility thesis when the researchers discovered the social gradient of disease independent of the access to medical and health services (12). The study that began in 1967 found that civil servants of the lower echelons of the British Civil Service had higher rates of deaths from coronary heart diseases than their counterparts in the higher echelons. The 25-year follow-up of the first *Whitehall* study participants confirmed the original discovery of an inverse gradient between mortality and the bureaucratic rank: the higher the rank, the lower the mortality. The social gradient in mortality was observed for most of the major causes of death. The study disclosed an important connection between the *distribution* of wealth and mortality rates. It is no longer the *absolute* income that is connected to mortality in developed countries. Rather, it is the *relative* deprivation, or the unequal distribution of

wealth that is connected to mortality rates. As people go down the social ladder, their level of income declines, although the people of lower social echelons are still above the poverty line. Thus, the real effect of income hierarchy is not associated with material deprivation, but it is the decline of sense of self-worth in the social environment that affects their health.

The relative deprivation manifests as *social status*, which affects health and illness through psychological pathways. As Richard Wilkinson argued, the *absolute* income affects morbidity and mortality of a population up to a certain income threshold that is sufficient to meet a certain standard of living (according to Wilkinson, Western Europe reached this income threshold at around \$5,000 GDPpc. at the 1995 price level). Once this income threshold is exceeded, it is the relative income that matters for mortality. Wilkinson concluded that psychological factors associated with social status, social relations and social interactions are critical intermediaries between socio-economic factors and disease (13). These intermediaries do not determine which specific disease will people get, or become vulnerable to. Rather they create a general susceptibility to diseases in the overall epidemiological context. If the majority of the population is impoverished, and the mean blood cholesterol level is low, no amount of stress could lead to an epidemic of heart disease. However, if the majority of the population is well nourished and the average plasma cholesterol level is high, the increase of stress could lead to heart diseases. The sense of lack of control and the feeling of alienation at the workplace and in society are related to the social gradient in cardiovascular disease. The sense of control means power and social status in social circles, and those who feel that they have no control over what is happening around them are often inclined to gain that control by turning to alcohol or violence, and are often unable to relax without a smoke. These habits, in turn, become the mediating factors between their social class status and morbidity.

Conclusion

Our understanding of what makes us sick, and what keeps us healthy inevitably contributes to what measures we take to prevent us from getting sick and

promoting health. The educational system, health care system, social organization and welfare policies in general reflect what are believed to be the major determinants of health and illness. The germ theory of disease that emerged in the late nineteenth century acquired recognition as epidemiologists responded to major infectious diseases around the world by successfully diagnosing the specific causative agents and developing treatments to cure them. The processes of diagnosis and treatment were further compressed by developing a classification of infectious and parasitic diseases as air-borne, water-borne, food-borne and vector-borne, which facilitated the actions against the major diseases. Thus, it contributed to a rapid succession of major discoveries in bacteriology, immunology and cellular pathology that led to the development of the biomedical model. These discoveries in biomedicine, together with the rapidly improving social and living conditions that reduced the hostility of the environment and increased the host resistance to disease, led to the decline of parasitic and infectious diseases since the mid-nineteenth century.

The multi-causal theory of disease emerged in the second half of the twentieth century with the epidemiological transition from predominantly infectious disease to degenerative diseases as the leading causes of death. The multi-causal epidemiology involved not only a major shift in the object of study and the recognition of the role of multiple causes, but also new techniques of studies and data analysis. Many of the contemporary health problems cannot be explained in terms of the germ theory of disease, as they are not directly connected to specific disease agents. Unlike infectious diseases, the complex development process of chronic diseases, such as coronary heart disease or cancer, often take a long period for the physical manifestation of the illness as specific behavioral, social, psychological and environmental factors themselves take shape as powerful determining factors in the behaviour of the individual. These factors eventually determine the level of disease in the community as a whole. In this context, the multi-causal analysis appears to be *holistic* in that it recognizes the importance of a multidisciplinary approach, and treats the population group as the unit of analysis.

The general susceptibility approach is not simply concerned with specific causal factors, but seeks to understand why certain population groups, such as unmarried, socially isolated communities, ethnic minorities, occupational groups and marginalized people in the population, are more susceptible to diseases than others. The general susceptibility approach could therefore be complementary to both germ theory and the multi-causal model of disease causation in health promotion and disease prevention strategies at the population level given that it not only takes into account the multiple causes of disease, but also examines their distributions at sub-population levels. If the majority of people today are dying because of chronic diseases, how useful is the germ theory of disease to prevent people developing heart diseases? Moreover, if divorced and widowed people are found to be more susceptible to high mortality rates from many different diseases, than the rest of the population, then how useful is it to know all the multiple causes responsible for the disease unless the curative and preventive measures are directed at these high risk groups? The general susceptibility approach essentially completes the methodological direction that the multi-causal analysis has initiated. As Mervyn Susser suggested, epidemiology must be broadened at its base, and move beyond its focus on individual risk factors to a new multilevel eco-epidemiology (14). Health does not exist in a vacuum apart from people, and people create societies, where structure and organization influence their behaviour. It is in these social structures that major causes of disease and death reside. The social epidemiology that examines social determinants of disease and their social distribution must be multidisciplinary and be more social scientific (15). Epidemiology is just one approach by which the major determinants of health in a population can be addressed, and it should be completed by historical and sociological research, both qualitative and quantitative, on a wide range of behavioural and socio-cultural aspects that have significant implications for morbidity and mortality rates in a given population. Effective social policies, and educational programs can modify these socio-cultural factors.

References

1. Soma Hewa, 2002 "Physicians, the Medical Profession, and Medical Practice," In B. Singh, Bolaria and Harley D. Dickinson (Eds.) *Health, Illness and Health Care in Canada*, Toronto: Nelson, Thomson Learning; Soma Hewa and Robert W. Hetherington, "Specialists without spirit: Limitations of the mechanistic biomedical model" *Theoretical Medicine*, 1995, 16: 129-139; Sara Nettleton, 1995 *The Sociology of Health and Illness*, Cambridge: Polity Press.
2. Kelyvn Jones and Graham Moon, 1987 *Health, Disease and Society*, London: Routledge and Kegan Paul.
3. Friedrich Engels, [1845] 1987 *The Condition of the Working Class in England*, p. 127. Middlesex: Penguin Books.
4. Laurence Foss and Kenneth Rothenberg, 1987 *The Second Medical Revolution, from Biomedicine to Infomedicine*, Boston: New Science.
5. Thomas McKeown, 1976 *The Role of Medicine: Dream, Mirage or nemesis?* London: Nuffield Hospital Trust.
6. Wade H Frost, "How Much Control of Tuberculosis?," *American Journal Public Health and The Nation's Health*, 1937: 759-766; Joseph Berkson, "Mortality and Marital Status: Reflections on the derivation of etiology from statistics," *American Journal of Public Health*, 1962, **52**: 1318-1329.
7. Lisa F Berkman and Cassel Syme, "Social Networks, Host Resistance and Mortality: a nine-year follow-up study of Alameda County residents," *American Journal of Epidemiology*, 1979: **109**: 186-204.
8. Ichiro Kawachi and Bruce P. Kennedy, "Health and Social Cohesion: why care about income inequality?" *British Medical Journal*, 1997; **314**: 1037-40; Ichiro Kawachi, Bruce P. Kennedy, Kimberly Lochner and Deborah Prothrow-Stith, "Social Capital, Income Inequality and Mortality," *American Journal of Public Health*, 1997; **87**:1491-98.
9. Emile Durkheim, [1897] 1951 *Suicide*, New York: Free Press.
10. Robert D Putnam, 2000 *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon & Schuster, p.331.
11. Jacob M Najman, "Theories of Disease Causation and the Concept of a General Susceptibility: A Review," *Social Science and Medicine*, 1980, **14A**: 231-7; Lisa F. Berkman

- and Ichiro Kawachi, 2000 "A Historical Framework for Social Epidemiology," In Lisa Berkman and Ichiro Kawachi (Eds.) *Social Epidemiology*, New York: Oxford University Press.
12. Michael G Marmot and Shipley M J, "Do Socio-economic Differences in Mortality Persist After Retirement? 25 year follow up of civil servants from the first Whitehall study," *British Medical Journal*, 1996; **313**: 1177-80.
13. Richard G Wilkinson, 1996 *Unhealthy Societies: The Affliction of Inequality*, London: Routledge.
14. Mervyn Susser, "The Logic in Ecological: I. The Logic of Analysis," *American Journal of Public Health*, 1994, **84**: 825-9.
15. Neil Pearce, "Traditional Epidemiology, Modern Epidemiology, and Public Health," *American Journal of Public Health*, 1996, **86**: 678-83.