

## Chapter 3

# POTENTIALITIES AND LIMITATIONS OF MULTILEVEL ANALYSIS IN PUBLIC HEALTH AND EPIDEMIOLOGY<sup>1</sup>

**Opening.** Like demography, discussed in the previous chapter, epidemiology traces its origins back to the investigations of John Graunt in the seventeenth century. But the two disciplines soon diverged because of their different objectives: demography seeks to understand how populations evolve in time and space, by interconnecting the phenomena that determine their size and composition, whereas epidemiology tries to understand the history of public-health problems affecting those populations and to combat them. As a result, epidemiology has evolved differently from demography over the centuries. In particular, it has used new aggregation levels and posed new questions, which this contribution will now examine.

The starting point at the aggregate level is, however, identical. The “sanitary statistics” that prevailed up to the mid-nineteenth century—along with their paradigm, “miasma”—were effectively situated at the aggregate level, where the sanitary conditions of populations were seen as the main causes of morbidity and mortality in those populations—far more so than any specific diseases. This led to improvements in sewage systems, drinking-water supply, sanitation systems, and so on.

But microbiological discoveries, particularly by Pasteur, revealed the existence of agents responsible for the transmission of diseases and spelled the end of the “sanitary” era. They heralded the era of “infectious-disease epidemiology” with its new paradigm: “germ theory”. The goal now was not to detect the effects of sanitary conditions, but to discover disease-causing germs, to curb their spread by means of vaccines, and to cure their effects through chemotherapy and antibiotics. The theory’s operating level shifted to the germ, which scientists needed to identify and study in order to combat it. This approach prevailed until the mid-twentieth century. It led to the eradication of several diseases through vaccination and to the development of antibiotics—although economic development and social change, as well, played a significant role in the process.

During this period, however, when the main infectious agents were apparently identified, new chronic diseases—of unknown origin—took centre stage: peptic ulcer disease, coronary heart disease, and lung cancer. “Germ theory” was incapable of supplying the resources to battle them. The new “chronic-disease epidemiology”, with its conventional biomedical paradigm, was free to flourish. It operated at the individual level, generating studies on “case control” and on cohorts monitored over long periods. The multicausal nature of public-health problems was emphasised, and the studies revealed the various causes of chronic diseases, whose origins were initially a mystery.

Ana Diez Roux shows us that this paradigm, as well, has been challenged in the late twentieth century. We must now take a broader view of epidemiology, with the aim of integrating the different earlier approaches into a more general model. The model must accordingly take simultaneous account of social factors (sanitary, sociological, and psychological conditions) and biological factors (genes, viruses, germs, individual risk factors, etc.). Moreover, these factors must be defined at multiple aggregation levels in order to incorporate advances in epidemiology: molecule, gene, virus, or germ level, individual level, social level, etc.

Ana Diez Roux examines how an epidemiology that tries to transcend the dichotomies in the “social/biological” and “groups/individuals” categories can help us understand not only the whole but also the parts—not only populations but the individuals in them. We are indeed witnessing the establishment of a paradigm of complex, hierarchical systems that may contain non-linear relationships. In these conditions, we may ask if the multilevel approach fully satisfies these expectations, or whether it is merely an analytical strategy providing a partial response to the need to examine factors in an overall context.

This chapter examines the possible reasons for using multilevel analysis in epidemiology, and discusses the theoretical and methodological issues raised by that use. What remedies does it

offer for the fallacies generated by the previous approaches: ecological fallacy, atomistic fallacy, psychologistic fallacy, and sociologistic fallacy? How does it integrate the aggregation levels described earlier? Does it allow us to develop more satisfactory strategies than the classical methods for analysing complex, dynamic systems, in interaction with their past? The answers to all these questions will give us a clearer view of the goals, potential, and limits of multilevel analysis in epidemiology.

D. C.

## 1. INTRODUCTION

Throughout history different types of factors have been emphasised as potential "causes" of disease (Susser and Susser, 1996a; Catalano, 1979; Tesh, 1990). In the early 19th century, disease was presumed to arise from foul emanations from soil, water and air (the miasmatic theory of disease causation), and emphasis was placed on the role of broadly defined ecological and environmental factors in disease aetiology. With the advent of the germ theory and the associated unicausal theory of disease causation, infectious organisms became the relevant "environmental" factors. Other aspects of the environment were important only to the extent that they were conducive to the reproduction or transmission of the "true biological cause". In the mid 20th century, the growing importance of chronic diseases led to the search for new causal factors, and research focused on the behavioural and biological characteristics of individuals as risk factors for disease. The study of the causes of disease thus shifted from the environment as a whole, to specific factors within the environment (biological organisms) and finally to the behaviours and biological characteristics of individuals. Concomitantly, the model of disease causation changed from a rather vague holistic determination, to the unicausal model of the germ theory, and finally to the multicausal model (the "web of causation") largely dominant in epidemiology today in which a variety of individual level biological and behavioural risk factors are presumed to interact in the causation of specific diseases (Krieger, 1994).

Over the past few years there has been an ongoing debate within epidemiology regarding the future of the field (Vandenbroucke, 1990; Krieger, 1994; Taubes, 1995; Charlton, 1996; Pearce, 1996; Schwartz *et al.*, 1999; Susser and Susser, 1996b; Susser, 1998; Poole and Rothman, 1998; Diez Roux, 1999; McMichael, 1999). Discussion of the origins and limitations the current epidemiological approach have been a key part of this debate. Much of the debate has centred around the critique of what is perceived to be the dominant paradigm in epidemiology over the past few decades: so-called "risk factor epidemiology", i.e. an epidemiology centred around identifying the "independent" contributions of individual level factors (usually biomedical or behavioural factors) to disease risk. The critique of risk factor epidemiology has had two distinct origins. On one hand it has been argued that epidemiology has severed its connections to biological and clinical science and hence has lost its ability to completely understand the causes of disease. On the other

hand, it has been argued that risk factor epidemiology has divorced disease from its social origins, reducing the causes of disease to biological or behavioural characteristics and isolating individual level characteristics (e.g. health-related behaviours) from the social contexts which promote their development and maintenance.

A key element of the current discussion pertains to the ways in which populations and population level factors are conceptualised and incorporated into epidemiological models of disease causation and empirical analyses. The concept of "populations" (as opposed to "individuals") has been key in epidemiology throughout its history. Indeed, epidemiology is often defined as the study of "patterns of disease occurrence in human populations and the factors that influence these patterns" (Lilienfeld and Stolley, 1994). Its object of study, "disease in populations" is often thus distinguished from the object of study of clinical medicine: "disease in individuals". Despite these general statements, however, the concrete ways in which the "population dimension" should be included in epidemiological models and empirical analyses has been rarely articulated or developed in mainstream epidemiology. In fact, the enormous methodological development and sophistication of epidemiology over the past decades has been marked by increasing "individualization", the notion that the risk of disease depends exclusively on individual level characteristics. This notion has been reflected in the behavioural model of disease (in which disease stems from the choices and behaviours of individuals, isolated from their social contexts), and reappears today in the genetic model (in which disease is strongly influenced by an individual's unique genetic makeup). In fact, much of today's epidemiology conceptualises populations merely as aggregates of individuals (useful from a statistical point of view), rather than as groups of interacting individuals with social relationships and social organisations, and with group level properties that may partly influence risk of disease (Pearce, 1996; Loomis and Wing, 1990; Almeida, 1992).

This process of "individualisation" of disease risk (i.e. attributing disease causation to characteristics of individuals rather than to environmental or social influences affecting populations) has also had methodological correlates. The underlying assumption that all disease determinants are best conceptualised (and consequently best measured) at the individual level has been accompanied by an emphasis on study designs where the units of analysis are individuals and where much effort is invested in the measurement of individual level characteristics, particularly behavioural and biological factors. Group level variables are used only as (often unsatisfactory) proxies for individual level data when the latter are unavailable. Interest centres in the examination of inter-individual variability and the individual level factors associated with it. Group-to-group variability (and the factors associated with it) is of little interest per se, except to the extent that it can be used to draw inferences regarding inter-individual variability. Thus, although there has been abundant discussion in the epidemiological literature of the fallacy inherent in using group level associations (ecological studies) to draw inferences regarding individual level relationships (i.e. the ecological fallacy) (Morgenstern, 1982, 1995; Piantadosi *et al.*, 1988; Greenland, 1992), there

Methodology and Epistemology of Multilevel Analysis

Approaches from Different Social Sciences

Courgeau, D. (Ed.)

2003, XII, 238 p., Hardcover

ISBN: 978-1-4020-1475-8