Invited Commentary

Invited Commentary: Genetic Variants and Individual- and Societal-Level Risk Factors

Steven S. Coughlin*

* Correspondence to Dr. Steven S. Coughlin, Environmental Epidemiology Service (135), Office of Public Health and Environmental Hazards, Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 22207 (e-mail: steven.coughlin@va.gov).

Initially submitted September 21, 2009; accepted for publication September 30, 2009.

Over the past decade, leading epidemiologists have noted the importance of social factors in studying and understanding the distribution and determinants of disease in human populations; but to what extent are epidemiologic studies integrating genetic information and other biologic variables with information about individual-level risk factors and group-level or societal factors related to the broader residential, behavioral, or cultural context? There remains a need to consider ways to integrate genetic information with social and contextual information in epidemiologic studies, partly to combat the overemphasis on the importance of genetic factors as determinants of disease in human populations. Even in genome-wide association studies of coronary heart disease and other common complex diseases, only a small proportion of heritability is explained by the genetic variants identified to date. It is possible that familial clustering due to genetic factors has been overestimated and that important environmental or social influences (acting alone or in combination with genetic variants) have been overlooked. The accompanying article by Bressler et al. (Am J Epidemiol. 2010;171(1):14–23) highlights some of these important issues.

coronary disease; genetic variation; genetics; genomics; heart diseases; myocardial infarction; polymorphism, single nucleotide

Abbreviations: CHD, coronary heart disease; HDL, high density lipoprotein; LDL, low density lipoprotein.

Over the past decade, several authors have called for new research paradigms in epidemiology that more adequately take into account different levels of analysis at the molecular, individual, and societal or group levels (1, 2). For example, in a far-ranging commentary published in the American Journal of Epidemiology in 1998, Diez-Roux (1) noted the resurgence of interest in the social origins of disease, echoing observations made by other epidemiologists about the importance of social factors in studying and understanding the distribution and determinants of disease in human populations (2–5). Diez-Roux noted that epidemiology may be on the brink of a new genetic paradigm and pointed to the development of new genetic technologies and the explosive growth in research on the genetics and molecular mechanisms of disease (1). However, the major thrust of her commentary was to underscore the importance of assessing both population risk factors for disease that exist at the group level (e.g., neighborhood effects) and more proximal risk factors that exist at the individual level (e.g., biologic variables or genetic traits assessed using the methods of molecular genetics). This approach allows biologic phenomena to be viewed within their social contexts and for individual-level explanations of disease causation to be integrated into broader models that incorporate interactions between individuals as well as group-level determinants and effect modifiers (1).

The approach outlined by Diez-Roux is consistent with the eco-epidemiology paradigm proposed by Susser and Susser (2), which encompasses many levels of organization, including molecular, individual, and societal. The hierarchical or multilevel structure of this model of disease causation extends from individual-level biologic or non-biologic risk factors to more distal societal factors. From
this perspective, incorporating information about the societal determinants of disease is key, since the model may overemphasize the importance of more proximal factors at the individual level if group-level or societal determinants of health are ignored (1).

In view of the calls by Diez-Roux, Susser and Susser, and others more than a decade ago for a new paradigm for epidemiologic research, it is worthwhile to consider the extent to which epidemiologic studies are integrating genetic information (and other biologic variables) with information about individual-level risk factors and group-level or societal factors related to the broader residential, behavioral, or cultural context. A recent commentary by Khoury and Wacholder (6), who defined “environment” broadly, noted that the proportion of articles reporting on gene-environment interactions remains at about 14% in the Human Genome Epidemiology Network literature. However, the proportion of articles that examined effect modification by group-level or societal determinants of health is likely to have been much smaller than 14%.

In this issue of the Journal, Bressler et al. (7) report that 2 of the genetic variants previously identified in Wellcome Trust Case Control Consortium case-control studies (rs1333049 and rs501120) were independently associated with incident coronary heart disease (CHD) among white participants in the Atherosclerosis Risk in Communities Study, even after adjustment for multiple established risk factors for CHD. The established CHD risk factors adjusted for by Bressler et al. (7) consisted of age, body mass index, high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol, diabetes, hypertension, and smoking. Other established risk factors for CHD (e.g., physical inactivity) and nontraditional or emerging risk factors such as prothrombotic factors were not controlled for in the analysis. Physical activity has been inversely associated with CHD in several epidemiologic studies after adjustment for other CHD risk factors (8). Possible pathways for a beneficial effect of exercise on CHD risk include the lowering of blood pressure and improvement of body composition, glucose tolerance, HDL cholesterol function, or thrombotic function (8). The relation between a sport activity index and CHD risk has previously been examined among participants in the Atherosclerosis Risk in Communities Study (9).

The decisions made by Bressler et al. about which CHD risk factors to adjust for in the analysis are important, partly because they considered a 2-sided $P$ value of <0.05 to be statistically significant and the magnitudes of the observed associations were quite modest. After adjustment for age, gender, body mass index, smoking, diabetes, hypertension, HDL cholesterol, and LDL cholesterol, the observed $P$ value for rs501120 was 0.01 and the point estimate of the hazard rate ratio was only 1.18 (95% confidence interval: 1.05, 1.33) (7). It would be helpful to have more information about why the authors preferred to use this particular set of risk factors for CHD and what procedures they followed to identify potentially confounding variables and effect modifiers, either at the individual level or the group level. Perhaps their findings might have been different if they had also adjusted for physical activity, socioeconomic status, alcohol consumption, or factors associated with thrombotic function. Of course, the pathways by which CHD risk factors exert their effects on CHD risk at the individual level (e.g., physical activity) or group level (e.g., neighborhood characteristics linked to alcohol consumption, diet, cigarette smoking, and obesity) may involve factors already addressed in the study by Bressler et al., such as HDL and LDL cholesterol or hypertension.

One of the issues raised by Bressler et al. is how to distinguish between “established” and “nonestablished” risk factors for CHD. Some frequently cited risk factors for CHD (e.g., obesity and body shape or waist-to-hip ratio) have been considered established risk factors by some authors but not others. Thus, in the context of evolving scientific information and diversity of opinions about what constitutes an established risk factor for common complex diseases such as CHD, it is likely to be important for researchers conducting molecular epidemiology studies to go beyond adjusting for a minimum set of traditional risk factors and to more thoroughly explore potential confounding and effect modification by several traditional and nontraditional risk factors.

What about the need to control for group-level variables, such as neighborhood effects, or to consider the possibility of effect modification by group-level variables? Should group-level variables be routinely incorporated into molecular epidemiology research such as the study by Bressler et al.? Although studies have demonstrated that group-level variables such as neighborhood characteristics influence CHD risk factors such as cigarette smoking, obesity, and socioeconomic status (10, 11), investigators studying group-level factors associated with CHD risk often test specific hypotheses and consider how neighborhood effects may be mediated through traditional CHD risk factors. In addition, the number of group-level variables that could be assessed in any given study (given sufficient time and resources) is very large. Testing for effect modification by a large number of group-level variables could result in some interaction terms being found to be statistically significant by chance alone, especially in studies involving large numbers of research participants. Thus, it is not clear that epidemiologists should routinely assess uncontrolled confounding or effect modification by group-level variables in studies that examine associations with individual-level variables, including genetic variants and other biologic factors.

Nevertheless, molecular epidemiology studies that do assess interactions with group-level variables may be informative, such as when they are designed to test specific hypotheses about possible causal pathways. Thus, there remains a need to consider ways to integrate genetic information with social and contextual information in epidemiologic studies, partly to combat the overemphasis on the importance of genetic factors as determinants of disease in human populations. As Ioannidis et al. noted, “There is a special enthusiasm about the potential power of genomics to define the etiology of disease and phenotypes” (12, p. 2). There may be a tendency to view associations with genetic factors as causal and associations

Am J Epidemiol 2010;171:24–26
with factors related to the broader social, cultural, or behavioral context as noncausal. Genetic and biologic factors are commonly considered to be foundational in hierarchical models and theories of disease causation, even though, as Parascandola and Weed (13) noted, there is no reason to assume that causes at the genetic or molecular level are any more “real” or significant than causes at another level, such as social factors. Even in genome-wide association studies of CHD and other common complex diseases, only a small proportion of heritability is explained by the genetic variants identified to date (14). It is possible that familial clustering due to genetic factors has been overestimated and that important environmental or social influences (acting alone or in combination with genetic variants) have been overlooked. There may be a tendency to inadequately take into account the complexity of the human biologic, physical, and social environment or the potential for confounding or effect modification by unmeasured genetic or environmental factors.

A particular challenge is the current lack of information about mechanisms of disease at the molecular, cellular, individual, and societal levels. Although an increasing number of genetic markers are being identified for CHD and many other diseases and health conditions, important questions remain about biologic and physiologic processes that may account for replicable associations with genetic markers. Genetic markers such as those examined in genome-wide association studies and some molecular epidemiology studies may be far removed from complex physiologic processes that are more important risk factors for disease (6). In addition, individual- and societal-level risk factors, examined in other epidemiologic studies, may represent only crude markers of disease risk and not provide detailed information about causal pathways.

In conclusion, although there has been increasing recognition of the desirability of not conceptualizing determinants of disease solely in terms of molecular or biologic factors and striving to also conceptualize them in terms of the broader social, cultural, or behavioral context, not all epidemiologic studies include or should include group-level or societal-level factors. For studies that do include such variables, investigators should consider developing a conceptual framework (i.e., a logic model) that clarifies the pathways by which various individual- or group-level variables, alone or in combination, are associated with the outcome of interest (3, 5).

ACKNOWLEDGMENTS

Author affiliation: Environmental Epidemiology Service, Office of Public Health and Environmental Hazards, Department of Veterans Affairs, Washington, DC.
Conflict of interest: none declared.

REFERENCES