The Right Answer for the Wrong Question: Consequences of Type III Error for Public Health Research

Sharon Schwartz, PhD, and Kenneth M. Carpenter, PhD

ABSTRACT

Objectives. This study examined the impact of assessing the causes of interindividual variation within a population when the research question of interest is about causes of differences between populations or time periods. This discrepancy between the research focus and the research question is referred to as a type III error, one that provides the right answer for the wrong question.

Methods. Homelessness, obesity, and infant mortality were used to illustrate different consequences of type III errors. These different consequences depend on the relationships between the causes of within- and between-group variation.

Conclusions. The causes of interindividual variation and the causes of variation between populations and time periods may be distinct. The problem of examining invariant causes deserves attention. (Am J Public Health. 1999; 89:1175–1180)

The goal of public health research is to provide a scientific basis for the development of effective strategies to improve the health status of the population. This involves investigating the causes of perceived threats to the public’s health. These causes may be distal or proximal, a characteristic of a person (e.g., weight), or a factor defined at a level of organization higher (e.g., poor neighborhood) or lower (e.g., genetic mutation) than the individual.

Since an infinite number of threats and causes could be investigated, decisions are made to limit the realm of inquiry. These decisions are determined not only by scientific concerns but also by social, political, and economic considerations. These value-laden considerations restrict the knowledge base available for public health strategies. This limitation has been extensively discussed. We argue here that additional value-laden constraints arise from research methods themselves because methods are designed to examine certain types of problems that, while critical to study, may nonetheless lead to restricted types of answers.

In lieu of the broad question “What causes this disease?” we typically investigate a more narrow set of questions often focused on what is different, changing, and unique. For example, why do some people in this population have the disease and not others? Why is the rate higher in group A than in group B? Why is the rate of disease increasing in country A? Problems sometimes arise because current epidemiologic methods are most developed for examining the first of these questions; that is, the methods are best suited for understanding the causes that distinguish individuals within a population. However, the questions posed are frequently not about interindividual differences within a population but rather about the cause of an increase in a disease rate or differences in the rate of disease between populations.

The use of methods addressing interindividual differences to answer such questions may be problematic because the causes of interindividual variation in disease within a population, causes of variations in the rate between groups and over time, and causes of the existence of the disease itself can all be distinct. Therefore, there is sometimes a discrepancy between the question being asked and the methods used to address the question. When the causes of the rate differences between populations (or time periods) and the causes of interindividual variation within a population are different and the question is about rate differences, this discrepancy results in what has been called a type III error—providing a right answer for the wrong question. In this report we discuss (1) types of relationships between causes of interindividual variation and causes of rate differences, (2) how a failure to recognize these distinctions can lead to type III errors, and (3) the consequences of type III errors for the goal of improving the public’s health.

Relationship Between Causes of Interindividual Variation and Rate Differences

A useful starting point for understanding why the causes of differences between individuals and between groups can be distinct is Rothman’s heuristic model for understanding causal inference in epidemiology (see

The authors are with the Joseph L. Mailman School of Public Health, Columbia University, New York, NY. Kenneth M. Carpenter is also with the New York State Psychiatric Institute, New York City. Requests for reprints should be sent to Sharon Schwartz, PhD, Columbia University, SPH 630 W 168th St, PHB332, New York, NY 10032 (e-mail: ssh5@columbia.edu). This paper was accepted March 2, 1999.

Editor’s Note. Please see related commentary by Koopman and Lynch (p 1170) in this issue.
Figure 1). The basis for the model is the recognition that causes work in conjunction to produce diseases. Sufficient causes are conceptualized as a combination of different component causes that act in concert. For any particular disease, there may be different combinations of component causes leading to many different sufficient causes, as illustrated in Figure 1, in which stroke is used as a hypothetical example. As Rothman notes, the effect size for any component cause is dependent on the prevalence of the other components necessary to complete this sufficient cause. For example, in sufficient cause 1, the effect size for diabetes would depend on the prevalence and joint distribution of both chronic stress and a genetic susceptibility to stroke.

The higher the population prevalences of chronic stress and genetic susceptibility (i.e., the complementary component causes), the higher the relative risk for diabetes, since more people with diabetes will have the other necessary components for the development of stroke (all things being equal). Therefore, in 2 populations with equal prevalences of diabetes and with equal prevalences of genetic susceptibility but different population prevalences of chronic stress, the strength of the association between diabetes and stroke (e.g., relative risk, risk difference, odds ratio) will be different. To take an extreme example, in the absence of chronic stress, the relative risk for diabetes will be 1. Therefore, any estimate of the relationship between a risk factor and a disease is dependent on the prevalence and patterning of the risk factors in the population.

While Rothman’s model is useful for illustrating the population specificity of the effect of risk factors, the model is not as useful for understanding the causes of these population differences in the prevalence and patterning of risk factors. We can build on this model, however, to examine the factors that may influence the differences in the occurrence of stroke in 2 different populations (see Figure 2). Assume that there has been a global economic crisis. In the first society (population A), with a history of economic inequality and norms emphasizing individual responsibility and initiative, the ensuing tight economy may lead to severe social competition. In the second society (population B), with an egalitarian ideology and norms stressing community values, the economic problems may not lead to any social competition but rather to increased feelings of community. In the first society, then, chronic stress becomes ubiquitous; in the second, it is absent. We will further assume, to simplify the example, that the 2 societies are identical in terms of the prevalence and patterning of all other components of all other sufficient causes.

If one did a study of people within population A, diabetes would be a potent risk factor for stroke, as would genetic vulnerability. Chronic stress would not be detectable as a risk factor for stroke in this population since it is ubiquitous; everyone is exposed to it, and it will therefore not differentiate people in this population who do and who do not develop stroke. One could detect differences in the recognition of and response to this stress, but not differences in the presence of the chronic stress itself, since this has become a normative characteristic of the society, something to which all members are exposed. We tend to talk about results from such studies as detecting causes of disease; more correctly, however, they can detect only a subset of causes, those for which there is interindividual variation within the population. Chronic stress is clearly a cause of stroke in this society, but it is not a cause of interindividual variation; that is, it is not a cause of why one person in this population has a stroke but another person does not. Within population B, chronic stress would not be a cause of stroke, since no one is exposed to it. Diabetes, however, would also not be a cause of stroke, since no one has the component cause—chronic stress—necessary for diabetes to have an effect on stroke.

Population A would also have a much higher incidence of stroke than population B. The cause of this rate difference is not due to differences in the prevalence of the causes of interindividual variation in disease identified within population A (i.e., diabetes and genetic vulnerability), since they are present in the same proportions and in the same patterns in the 2 populations. Rather, the rate difference between the populations is due to differences in egalitarian norms that lead to social competition and the ubiquity of this stressor in population A but not in population B. There has also been a change in the incidence of stroke in population A over time, owing to the global economic crisis resulting in the increase in exposure to chronic stress.

What this example illustrates is that the causes of interindividual variation in risk within a population (e.g., diabetes, genetic vulnerability) may be distinct from the causes of the differences in the disease rate over time (i.e., global economic crisis), which may be distinct from the cause of the rate difference between populations (i.e., differences in egalitarian norms). All of these factors, in turn, are distinct from the cause of the existence of the disease itself—an inability of the brain to exist without oxygen. What is the cause of stroke? Clearly, all of these factors. The problem of type III error can arise when methodologies designed to address the question of interindividual variation (differences between individuals within a population) are used to address all of these questions. We tend to overlook the fact that examining the distribution of a disease within a population only tells us about causes that can be distinguished among these individuals. This method cannot detect any causes of disease that are ubiquitous or relatively invariant within the population under study, whether they are inherently group characteristics or derive from interactions among individuals.

Different aspects of the distinction between the causes of a disease’s existence and its rate and/or rate differences over time and across places have been explicated in many fields, including behavioral genetics, biology, sociology, and public health. Despite this fact, researchers sometimes...
examine the causes of interindividual differences in risk within a population to answer questions about rate differences between populations. In epidemiology, this tendency is supported by the current paradigm's focus on individual-level risk factors. In addition, the necessity for variation also supports this tendency. Covariation has been identified as a necessary, if insufficient, indicator of causal relationships. Since all statistical methods require variation, research questions are best addressed with data that include substantial variation in the variables of interest. Because variation is often maximal at the individual level as opposed to the population level for factors most frequently examined in epidemiology, interindividual differences within populations often become the focus of attention, even when the original question is not about such interindividual differences.

What are the consequences of this type III error for the public health goal of reducing a population's incidence of disease? We think that the consequences depend on the relationship between the causes of the differences among individuals within a population and the causes of the rate differences (between populations or time periods) in their contribution to the current incidence in the population that is the focus for intervention. In this article we examine 3 areas of public health research where type III errors occur—homelessness, obesity, and ethnic differences in infant mortality. Each area was chosen to serve as a heuristic device for the explanation of a different relationship between the causes of interindividual differences within a population and the causes of rate differences between populations or time periods (see Table 1).

We argue that examining causes of interindividual differences in risk for homelessness is not useful for appreciably decreasing the incidence of homelessness, because the causes of interindividual variation in risk for homelessness do not appreciably contribute to the current incidence of homelessness. In obesity research, examining interindividual differences may have some public health usefulness, but of a limited nature, because the causes of the rate differences between populations and the causes of interindividual differences interact in creating the current incidence in the population of interest. In the case of ethnic differences in infant mortality, examining interindividual differences may have important public health consequences, but not the ones wanted, because while the causes of interindividual differences are also causes of the current incidence, they may not be causes of the rate difference between the groups of interest.

**Examples of the Problem:**

**Homelessness, Obesity, and Low Birthweight**

**Homelessness**

Immediately before the 1980s, homelessness as a social problem was almost nonexistent. This rate change over time motivated considerable research about the causes of homelessness. Most studies, including those in the public health literature, examined the characteristics that differentiate homeless from nonhomeless people. Thus, individual risk factors such as physical and mental health problems, childhood foster care, and drug and alcohol use have been identified as causes of homelessness.

Differences between people who are and are not homeless at any point in time may pertain to the question of who becomes homeless but not to the cause of the rise of homelessness over time. This is because the causal factor of interest must entail a change in some factor occurring before and after the time when homelessness emerged. In assessments of the characteristics of individuals after 1980, this change has already occurred and is therefore held constant. Structural factors such as the amount of affordable housing largely determine the rate of homelessness, whereas individual-level risk factors are likely to determine who becomes homeless. Yet even researchers who recognize that these are distinct issues with distinct causes often conclude that both types of studies are important for reducing the rate of homelessness. We argue, however, that homelessness represents a situation in which information about interindividual differences (i.e., individual-level risk factors) is not necessary for reducing the rate of homelessness because the causes of the current amount of
Table 1—Consequences of Type III Error for Decreasing an Outcome

<table>
<thead>
<tr>
<th>Contribution to Current Rate in Population of Interest</th>
<th>Consequence</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only causes of between-group or time period differences make contribution</td>
<td>Interindividual differences uninformative</td>
<td>Homelessness</td>
</tr>
<tr>
<td>Causes of between-group and interindividual differences relate</td>
<td>Interindividual differences informative but inefficient</td>
<td>Obesity</td>
</tr>
<tr>
<td>Each makes independent contribution</td>
<td>Interindividual differences interactive, but not necessarily for decreasing group differences</td>
<td>Black-White differences in low birthweight</td>
</tr>
</tbody>
</table>

Data (1997) concomitant with an increase in the occurrence of obesity. A large number of research papers, apparently motivated by this increasing health problem, have examined interindividual differences in body mass. Unfortunately, interindividual differences in obesity at any point in time are not likely to provide clues to the causes of this rate increase. For example, studies have shown that genetic factors account for a significant proportion of the variance in body mass within a population. However, it is extremely unlikely that genetic factors could explain the rate increase, since genetic variation does not change that rapidly.

Examining interindividual genetic variation at any point in time holds constant whatever other changes have occurred between time periods (e.g., an increase in the perversiveness of advertisements enticing people to eat, the number of fast food restaurants per square mile, or exposure to sedentary leisure activities). Genes per se do not cause obesity; rather, for example, they produce proteins related to a greater or lesser propensity for appetite dysregulation. The threshold at which a genetic predisposition to obesity is expressed is likely to be lower in an environment where food is plentiful than in an environment of limited food resources. Because the environmental factors are ubiquitous (those exposures do not vary), they will not explain any variation in the distribution of the disorder. If one confuses the reasons behind the distribution in this population with the causes of the increase in the rate in the population, very misleading conclusions may be reached about the causes of the rate increase in obesity.

Information about individual-level factors in such instances may have public health utility since the causes of interindividual differences within the population (e.g., genetic propensity for appetite dysregulation) interact with the causes of the rate increase (e.g., increased food advertising) to produce the current incidence in the population. Unlike the situation with homelessness, the environmental context does not necessitate that some individuals become obese. As we have seen, in any interaction, if 1 component cause is missing, the outcome is prevented. However, in situations in which a factor responsible for the increase continues to rise while the causes of the interindividual differences are stable, interventions based on the latter are likely to be ineffective. If the cause of the rate increase is left unaddressed, then an increasing number of individuals will fall into the “vulnerable” category over time as the contextual factors become more widespread and/or potent. In addition, such interventions place the onus on individuals to adapt to increasingly high-risk environments (e.g., inducements to eat unhealthy foods in unhealthy amounts, promotion of sedentary activities).

Black-White Differences in Infant Mortality:

The insouciant Black-White discrepancy in infant mortality is another public health issue that has commanded considerable research attention. The infant mortality rate among Blacks is about twice that among Whites, even within socioeconomic strata. Some researchers have suggested that this question be addressed through an analysis of risk factors for preterm delivery among Black women. While this may be useful, it may not answer the question of interest, which is “What are the causes of the rate difference between Blacks and Whites?” In restricting the analyses to Black women, any risk factor that is ubiquitous for Black women (e.g., exposure to discrimination) is held constant and therefore will not be identifiable as a risk factor. Such a design can only uncover factors that distinguish which Black women are at greater or lesser risk of losing an infant (e.g., viral infections, nutrition, recognition of and response to discrimination). The variables found to account for risk differences among Black women may not provide the correct answer for explaining group differences.

August 1999, Vol. 89, No. 8
Studies that investigate interindividual differences in risk for infant mortality among Black women, when the question concerns rate differences between Blacks and Whites, represent our third case scenario. In this situation, information about interindividual differences is likely to be useful in decreasing the rate of the outcome, because the causes of the intragroup variation are likely to make a distinct contribution to the infant mortality rate. The individual-level risk factors may indeed be causes of infant mortality, and interventions based on them will likely be successful in reducing the rate. Such interventions will not change the size of the group differences, however, unless the group difference was just a reflection of differences in the prevalence of the causes of interindividual differences within each group. In any case, this could not be verified without appropriate comparisons between groups. If the causes of the group differences and the intragroup variation are distinct, interventions based on individual-level factors in both groups may reduce rates among both Blacks and Whites but will not necessarily reduce the rate difference between Blacks and Whites if there is a contextual or structural factor associated with it. Whether a focus on interindividual differences or on group differences is more effective from a public health standpoint (i.e., from the point of view of decreasing the rate) depends on the rate of disease in each group and the magnitude of the group differences.

Conclusions

When the question of interest is about risk differences between groups or time periods, the answer requires examination of multiple groups or multiple time periods; otherwise, a type III error can result. The assumption underlying the concept of the type III error is that looking at inter-individual differences is an error or that group differences can never be a reflection of inter-individual differences; rather, it is that risk differences between individuals within a particular population may not have the same causes as the differences in the average risk between 2 different populations. We examined 3 specific areas of public health research in which this problem occurs. They were chosen to serve as heuristics, and they do not represent the only areas in which these problems arise. In addition, the examples we used do not represent the full range of possible relationships among the risk factors for obesity and low birthweight. Nonetheless, we think they illustrate different consequences of examining the causes of interindividual variation within a population when interest is in rate differences between populations or time periods.

The focus on interindividual variation has value-laden and political implications because such analyses, implicitly or explicitly, consider ubiquitous exposures uninteresting, unchanged, or outside the purview of epidemiologic consideration. If the effects of such exposures and contexts are not investigated, they are not as available for intervention. To examine such exposures requires their overt consideration and different sampling, measurement, and conceptual frameworks.

Research with potentially useful public health consequences requires consideration of the full range of risk factors at all levels of organization. At a minimum, restricting our focus to a particular level of organization provides a narrow knowledge base for intervention; at the extreme, type III errors can lead to research with little potential for significant public health consequences (e.g., research on interindividual differences in risk factors for homelessness provides an inadequate basis for interventions designed to significantly decrease the rate of homelessness). In either case, we need to rethink the issue of levels of organization and take seriously the problem of ubiquity—that is, the difficulty in detecting invariant causes. Causes that are invariant within populations or historical moments are precisely those that determine which interindividual differences are important and are likely to affect a population's health. An exclusive focus on interindividual differences leaves unanalyzed this important and consequential class of etiologic factors.

Contributors

This paper represents a full collaboration between the 2 authors. S. Schwartz conceived of the original idea, which was developed during numerous meetings with K. M. Carpenter. S. Schwartz drafted the first rough version of the paper, which was extensively revised during meetings with K. M. Carpenter. Both authors take equal responsibility for the contents of the paper.

Acknowledgments

This research was supported in part by National Institutes of Mental Health grant MH 13043.

We would like to thank the following people for their comments on earlier drafts of this paper: David Allison, Richard Blumenthal, Myles Faith, Paul Landsbergis, Ian Meyer, Ann Stueve, Mary Beth Terry, J. Blake Turner, and the anonymous reviewers.

References


American Journal of Public Health