ETHNICITY AS A TAXONOMIC TOOL IN BIOMEDICAL AND BIOSOCIAL RESEARCH

DOUGLAS E. CREWS, PhD, AND JAMES R. BINDON, PhD

Ethnicity and race are commonly used in biomedical research to portion human populations into groups for analysis. The view taken here is that race is a sociological construct that is poorly correlated with any measurable biological or cultural phenomenon other than the amount of melanin in an individual's skin. On the other hand, ethnicity is a sociocultural construct that is often, if not always, coextensive with discernible features of a group of individuals. These features include, but need not be limited to, language, style of dress and adornment, religion, patterns of social interaction, and food habits. The purpose of this brief commentary is to examine definitional issues surrounding use of the term ethnicity, to review the use of ethnicity in biomedical and biosocial research, and to examine the ecological and evolutionary basis for ethnic differentiation in the human population. Its purpose is not to provide a final definition of ethnicity, but instead to suggest that ethnicity is a multiplex concept that has a variety of applications and definitions, each perhaps dependent upon particular research problems and situations. We suggest that individual researchers explicitly state how they are using ethnicity, what their chosen categories imply biologically and sociologically, and why their particular analyses are needed. This may help to limit conflicts that arise over the inappropriate use of ethnicity in biomedical and biocultural research. (Ethnicity Dis 1991;1:42-49)

KEY WORDS Ethnicity, Disease, Research Design

The finding that health and disease exhibit geographical and population variation is nothing new. Causes for differences in disease patterns across various groups of humans have intrigued physicians, epidemiologists, and philosophers for centuries." Environment and behavior have long been established as important factors for differences in disease etiology and outcomes. One term that has been used to categorize a number of these differences between peoples is ethnicity.

The majority of health researchers and practitioners have a concept of ethnicity, whether personal or disciplinary, that they apply when reading or hearing about ethnic differences. In research, ethnicity frequently is used to categorize human groups; however, few investigators have attempted to state their concepts explicitly. Because ethnicity is infrequently defined, its usage in specific circumstances may remain elusive. Often white and black are used to represent an ethnic taxonomy. Divisions based on skin color, however, represent unique problems, and racial categories so constructed commonly have little relationship to ethnicity. Each category is likely to be composed of several different ethnic groups. Although the use of race in the study of human biological variability may have applicability in particular settings, it is a poor reflection of social, cultural, or biological variability. There remains a need for a valid concept that will allow researchers across various disciplines to subdivide the human species consistently and reliably. Ethnicity is a concept that may fill this need.

From the Department of Preventive Medicine and Epidemiology, Loyola University Stritch School of Medicine, Maywood, Illinois (DEC), and the Department of Anthropology, University of Alabama, Tuscaloosa (JRB).

Reprint requests to Department of Preventive Medicine and Epidemiology, Loyola University Stritch School of Medicine, Maywood, IL 60153 (Dr Crews).

Received October 9, 1990; accepted in revised form November 1, 1990.
In this paper, we describe present usages of and inherent problems with *ethnicity*. Next, we examine some of the multifactorial aspects implicit in a valid concept of ethnicity and identify a number of factors that may be useful in operationalizing ethnicity in specific medical and epidemiological settings. Last, we suggest, as have others, that ethnicity is not a simple concept, and that, at least to some degree, the definition to be used depends on the nature and requirements of the specific research project. The perspective presented here relies heavily on concepts and theories that have been developed within the research area of human population biology. However, we feel that these concepts have broad applicability within biomedical and biosocial sciences, and that they may be useful to other investigators studying population, racial, or ethnic variability in health and disease.

**Previous Research and Problems**

One problem that has plagued much previous research in this area is the use of *race* as synonymous with *ethnicity*. Problems with racial categorizations have been widely documented and need not be reiterated here. It has been suggested that the genes responsible for morphological features that allow classification into broad racial groupings are atypical and not representative of the true degree of interracial genetic differences. It is also of interest to note, in this context, that cultural restrictions for mate selection based on external morphological characteristics may lead to rapid differentiation of skin, eye, or hair color between populations while leaving aspects of basic biology and energy metabolism relatively unaffected.

Racial categories are products of a typological classification. The classification is based on ideal types, one of which (black) is shown to differ in a specific manner from another (white). However, intermediate types that do not exactly match the ideal (eg, brown or almond) cause the typology to break down, because such variants must be either forced into one of the ideal classes or ignored. Ultimately, classes include specimens not identical to the original archetype. A major problem is that racial categories are not coextensive with any existing ethnic group, particularly if the term *race* is taken to presuppose cultural differences. For example, *white* includes English-speaking peoples of Great Britain and Croatian speakers of Yugoslavia; although they are of the same race, few would argue that they represent a single ethnic group. In studies of "US whites," however, historical migrants to the US from both Britain and Yugoslavia are necessarily included as white. A similar situation exists for "US blacks." Descendants of black West Africans, from numerous different tribal groups who came to the US in forced migrations, are combined with black East Africans, who migrated during different eras for various reasons, into a single class. Yet these groups may have less in common, other than a darker skin color, than those identified as white.

A different, but related, problem is that *ethnicity* is a vague term that has been loosely used in medical sociology, epidemiology, and anthropology. *Ethnicity* is an almost universal term in cosmopolitan medical records and health questionnaire research, yet it is a poorly defined categorization. In most cases in which ethnicity is reported, either no definition or only a superficial one is given, and categories are frequently both arbitrary and nonexclusive. That is, the same individual may fall into one or more classes of a particular scheme but must be forced arbitrarily into, or must choose, only one. This vagueness is also seen in the looseness of the standard dictionary definition: *ethnic, "2. designating or of any of the basic divisions or groups of mankind, as distinguished by customs, characteristics, language, etc."* Another definition is "... a member of a minority or nationality group that is part of a larger community," whereas
**Ethnicity** is defined as "ethnic classification or affiliation." These definitions, particularly the "etc" in the first one cited, permit a flexibility in interpretation infrequently encountered in scientific research and terminology. They also support the idea that ethnicity is a multiplex concept with a variety of possible contributing attributes.

**Multifactorial Aspects of Ethnicity**

Use of the term *ethnic group* to describe subdivisions of humanity has been vigorously advocated in the field of human biology by Ashley Montagu for over 50 years. He cites two antecedents who influenced his use of the term. The first was Dekiker, who says, "... ethnic groups are formed by virtue of community of language, religion, social institutions, etc., which have the power of uniting human beings of one or several species, races, or varieties, and are by no means zoological species. ..." Montagu acknowledges Huxley and Haddon as the immediate source from which he drew the term. They note that "The essential reality of the existing situation ... is not the hypothetical sub-species or races, but the mixed ethnic groups, which can never be genetically purified into their original components, or purged of the variability which they owe to past crossing. Most anthropological writings of the past, and many of the present, fail to take account of this fundamental fact." Montagu favors the use of *ethnic group* to explicitly question hypotheses about the genetic and cultural constituency of the group. He says, "... when one uses the term 'ethnic group,' the question is immediately raised, 'What does it mean? What does the user have in mind?' And this at once affords an opportunity to discuss the facts and explore the meaning and falsities enshrined in the word 'race' and to explain the problems involved and the facts of the genetic situation as we know them." It is an unfortunate fact that most researchers who have adopted the use of the term *ethnic group* have not attempted to understand the concept of ethnicity. Most still apply a genetic meaning, which equates *ethnic group* with *race*, neglecting the cultural component of ethnicity. That this problem is of long standing is indicated by Huxley and Haddon as they try to contrast the social context of *ethnic group* with the genetic significance of *race*: "If race is a scientific term it must have a genetic meaning."

Evidence of the misuse of *ethnic group* in the study of human biology is widespread. Heath and colleagues found an ethnic factor in the differences in somatotype between Japanese- and European-derived samples. They cite the ethnic factor as evidence of genetic control of the somatotype. Similarly, Hiemaux compared two Rwandan groups, the Tutsi and Hutu, and showed that the Tutsi were tall and linear, whereas the Hutu were short and stocky. Because he found that the nutritional content of their diets was similar, he concluded that the differences must be genetic. More recently, Clegh discussed differences in the growth of Melanesian and Indian children in Fiji. He noted dramatic differences in social, cultural, and economic niches between the two groups, but he attributed most differences in growth and maturation to racial (ie, genetic) causes. In each case, ethnic group differences are equated with genetic or biological differences at the expense of sociocultural explanations for the observed variability.

"The fact that many people still misuse the concept of ethnicity in research does not mean that it is a valueless notion. Some workers have attempted to answer the questions posed by *ethnicity*, but there is much room for refinement. An early proponent of untangling the ethnic question, Baker discussed the importance of simultaneously considering genetic and cultural dimensions when analyzing climatic adaptation. One of the better recent operationalizations of
Ethnicity as a Taxonomic Tool—Crews & Bindon

The population involves research on Mexican Americans. Hazuda and coworkers have established an algorithm to determine Hispanic ethnicity based on father’s surname and mother’s birth name, birthplace of parents, subjects’ preferred ethnic identity, and reported ethnicity of all four grandparents. They recognized separate ethnic groups within Hispanic whites (e.g., Cubans, Central Americans, Puerto Ricans, etc), and they acknowledged the sociocultural nature of the distinctions among these different groups. In addition to documenting gross differences in obesity and diabetes between non-Hispanic whites and Mexican Americans in their San Antonio project, Hazuda and coworkers proceeded to probe the underlying social causes of these differences and did not simply accept the difference as the result of genetics. Toward this end, their study employed sophisticated measures of socioeconomic status as well as several dimensions of acculturation.

Although these attempts to understand the cultural dimensions of ethnicity are necessary, they are not sufficient to answer Montagu’s questions about the ethnic group. For we must remember that ethnicity usually contains information about the history of the population, and hence the genetics of the group, as well as sociocultural information. As the San Antonio group has recently pointed out, among Mexican Americans in particular, the degree of admixture with Native American populations may play an important role in determining biological responses. Similarly, Weiss and colleagues have provided data on the increased risk of cholecystectomy, diabetes, and gallbladder cancer in Mexican Americans with increasing Native American admixture. Thus, human biological research focusing on ethnic differences must account for both genetic and sociocultural dimensions.

The Population

In human population biology, a fluid transdisciplinary construct based upon multiple criteria is used to specify the “population” for analysis. The population may be a small unit within a larger grouping (such as the Mennonite members of a single church in Nebraska), or an isolated or semi-isolated segment of a larger population (such as the Polynesian residents of the islands of American Samoa, or a heterogeneous group (such as the residents of Sun City, Arizona). In the first case, religious and sociocultural criteria, and in the second, genetic, geographic, and sociocultural criteria were used to define a population of Mennonites and American Samoans, respectively. In the third, however, the major unifying characteristics were age and place of residency. Thus, population, as used in human population biology, relies on multiple criteria specific to the research hypothesis. A similar type of conceptualization may aid in the development of an acceptable definition of ethnicity. In addition, the ecological, evolutionary, and cultural perspectives of human population biology may aid in developing a framework for multifactorial criteria to be included in a widely applicable definition of ethnicity.

In a construct similar to that of human population biology, numerous individual attributes may be useful for delimiting population subgroups. At least two possible dimensions, ecology and environment, suggest that adaptation to a similar niche at some point in time needs to be included as an aspect of subgroup definitions. Other dimensions include a common language, which is also suggestive of a historical relationship among individuals. Furthermore, language, unlike other factors, is easy to identify and remains relatively stable over time; although language groups are open to nonnative speakers, other criteria should exclude these individuals. Diet and patterns of food use can be related to religious, ecological/environmental, or cultural factors and may also be useful in defining ethnic subgroups. Some groups may attach sociocultural values to various foods. For
example, the use or nonuse of pork by certain ethnic groups may have an adaptive aspect associated with parasitism,26 but in other cases may be a religious belief associated with different sects. Sociocultural factors also include style of dress, nationality, and proximity. Style of dress often is a conscious attempt to maintain an ethnic identity. It appears that each of these factors may aid in defining an ethnic group. Any one of these might be a sufficient factor in a particular setting, but none may ever be a necessary condition for defining ethnicity. This may be the crux of the problem in establishing a universally acceptable concept of ethnicity: many criteria, including self-identification, are equally sufficient conditions for the task, but none are necessary conditions.

The conclusion may be that an ethnic taxonomy is unlike other disciplinary taxonomies. In an ethnic taxonomy, different criteria can lead to different but equally valid groupings. This may present an insurmountable barrier to those who have been imbued with the need for a strict classification system (such as that of Linnean species or the periodic table of elements), but for social and health scientists this demonstrates the wealth of comparative data available for research. However, even Linnean species are permeable at their edges, as fertile cross-matings between different species of primates, ungulates, and canines have shown. Conversely, the periodic table includes items that are mutually exclusive by virtue of their substructure; although the periodic table provides a model for classification, elements do not possess the attributes of living organisms.

This suggests that researchers in medical sociology, social epidemiology, and medical anthropology may, in defining populations or ethnic groups, use a variety of possible criteria that are valid for their particular research purposes. The real problem becomes one of adequately specifying the multifactorial grouping criteria used in each analysis rather than of debating what is or is not an ethnic group, a view recently proposed elsewhere.5,27 Recent developments in molecular biology suggest that, in the future, it may even be possible to specify variation at a microevolutionary level that relates to differences in ancestry, culture, language, and food habits of specific populations within an area that might otherwise be considered to encompass a single ethnic group.28-30 Once defined, these fine subdivisions may then be examined as separate ethnic groups with respect to biomedical risk factors for specific diseases or biosocial influences on risk factor levels.

The ecological-evolutionary perspective in health research is not new. It has been a guiding principle of medical anthropology and human population biology for many years.4,5,17,31-33 At its core is the observation that populations or ethnic groups evolve in response to the environment. These adaptations can be biological or cultural, or a combination of both (biocultural). Differences in cultural adaptations are likely to be more common between groups, because they may occur more quickly in response to particular stresses. Eventually, the health status and medical belief systems of different groups come to differ because of the variety of possible ecological responses to environmental differences. Landy33 discussed how differences between social groups in industrial societies may lead to variable health care usage patterns. It is the interrelationships among adaptability, culture, and the evolution and geographic distribution of disease that researchers examining ethnicity and disease are attempting to document by their classifications of human groups.34 What this leads to is the ethnic group, its sociocultural pattern, and the associated ecosystem as the unit of analysis. Ethnic differences in physiology, genetics, beliefs, and lifestyle may then be viewed as the result of biological and cultural adaptations by particular human groups that have maintained some degree of relative cultural
and/or reproductive isolation for a period of generations in a particular environment. This implies that the group has existed under a particular set of conditions (climato-logical, ecological, social) for a sufficiently long period to be distinct from other similar groups in a myriad of factors—language, religion, genes, relative skin color, diet, behavior, habits of dress, and social interaction. These factors may therefore be used collectively or sometimes individually to define the ethnic group.

It is important to note that the major and minor geographic and cultural regions of the world are likely to be composed of a wide variety of ethnic groups that are more or less distinct from each other. The British Isles provide one illustration. Closely related populations—Irish, Scottish, and English—all share a common heritage, yet manifest their distinctiveness in social interaction patterns, religion, diet, and language. A different situation exists in the US, where descendants of these groups residing in the northeastern states appear more similar and often are grouped together as Anglos or white Anglo-Saxon Protestants. As another example, on the small peninsula of Peljesac, along the Dalmatian coast of Croatia, are a number of small rural villages whose inhabitants’ principal activities are farming and fishing. One might think that, because they are geographically close and well defined as part of the Dalmatian coast population, they may be treated as a single ethnic group for analyses of disease and related physiological and genetic risk factors. This, however, would be an oversimplification. At least three different population groups can be defined on this small peninsula. They are the products of adaptive responses that occurred during the previous 5 or 10 centuries and of microdifferentiation between adjacent valleys and villages. Each group continues behaviors, styles of dress, and dietary traditions received from founding ancestors who migrated to the peninsula from different places at different times. They thus show differences today in morphology, polymorphic blood protein frequencies, and isonomy. Such studies illustrate the possible extent of microdifferentiation in human groups. Furthermore, they suggest that variation in the purpose of the proposed research may lead to an ethnic group being loosely or very specifically defined, thereby leading to highly generalized or culturally specific research results.

Most academic researchers today are the product of pluralistic societies in which ethnicity is thought of in terms of nationality (eg, Italian, Greek, German). These macro-level constructs may have little validity outside large nation states, because they do not relate a population to its evolutionarily important microecological situation. They do, however, represent variability in language, sociocultural norms, diet, and belief systems and may be valid for examining some problems related to disease onset and progression. However, at a finer level of analysis, it becomes necessary to examine population groups with a closer association to their ecosystem. Samoans are all members of a single identifiable Polynesian group that, prior to European contact, generally resided on a small series of volcanic islands in the South Pacific. However, macrolevel analysis of the Samoan population must now include Samoans residing not only in the Samoan Islands, divided into American and Western Samoa, but also in Hawaii, on the US mainland, and in New Zealand. These groups have been exposed to variable stressors in recent decades, and to examine “true” Samoans would be impossible. However, the most isolated group in Western Samoa may provide a baseline from which to examine relationships between health and ethnicity among those more exposed to new life-style factors. Similarly, US Italian samples from New York City may reveal little about their counterparts in remote valleys of Italy but may illustrate a good deal about Italians residing in Rome. These examples reveal that traditionally defined ethnic
groups may today be adapting to a wide variety of environments through modifications of a similar biological and cultural heritage. It is the responsibility of the researcher to specify these conditions and their criteria for inclusion of individuals into an ethnic category. It is to be expected that such criteria will vary in keeping with the research problem and should no longer be limited to macrolevel socioculturally un-related groupings such as black/white/Anglo/Asian/Hispanic/other.

In trying to call attention to some of the shortcomings in the use of ethnicity in research, an analogy may prove useful. When gathering biomedical data, we are careful to consider all possible causes of measurement error. Thus, when reporting on blood pressure measurements, descriptions include the equipment used to make the determinations, the physiological indicators recorded (eg, the Korotkoff sounds), the number of determinations made, the posture of the subject, and the physical and behavioral context. In statistical analyses we test the assumptions of our model (normality, linearity, heteroscedasticity, etc); when untested analyses are presented, skepticism is warranted. However, the same study may present no discussion at all of how ethnicity was determined. It is necessary to exercise the same care used for biomedical measures and analyses if an understanding of the multiple dimensions and underlying assumptions of ethnicity is to be obtained. It is not sufficient to conclude that ethnic differences are genetically induced when no significant difference on some dimension of socioeconomic status is observed. Testing a genetic hypothesis requires genetic data, such as are derived from linkage studies or familial aggregation, and use of appropriate statistical techniques.

Baker argued for research designs that explicitly test the genetic and social components of ethnic variability. Under the rubric "one stress—multiple populations," he illustrated how comparisons can be carefully crafted to test for differences due to genetic as opposed to social factors. These may include, for example, migrant studies or comparisons of groups responding to different levels of a similar stress while remaining in their native environment. Good research on ethnicity and health must go beyond simple measurement and comparison of black/white/Anglo/Asian/Hispanic/other groups to the construction of research designs adequate to explore the multifactorial nature of ethnicity.

REFERENCES


