Contemporary European and American research on race, ethnicity, and health uses poorly defined labels to describe study populations. The search for accurate terminology remains controversial, for scientific and social reasons, as illustrated by discussions of the terms Hispanic and Asian and the changing meaning of ethnicity and race in the United Kingdom and the United States. 

Researchers should describe the study populations, define the terms used, and avoid lumping together heterogeneous populations.

The controversy and scholarly and political debate have largely bypassed the label White. Words such as White and Caucasian have been accepted as self-evidently suitable. While, for example, the labels Hispanic and South Asian are indexed in key textbooks, White is not. White is institutionalized in the censuses of the United States and the United Kingdom and is used freely in compiling statistics and epidemiology in clinical medicine.

Racial and ethnic nomenclature in the United States is dominated by the classification of the Office of Management and Budget, which was devised by a subcommittee of the Federal Interagency Committee on Education. It was adopted in 1977 and has been reviewed. Its purpose was to collect data on groups that are discriminated against, primarily for monitoring civil rights. The recent review drew little attention to the category White, although there were requests to include European-American, German-American, and Arab-American categories.

The US census has 10 subcategories for Asians and Pacific Islanders but none for Whites, who accounted for 80% of the population in 1990. The British census is similar. White is effectively a category for everyone left out of specific racial and ethnic groups. Researchers tend to rely on these classifications because of cost and convenience, even though they were devised for a different purpose other than research. As health and social research is an important by-product of statistics on population, this debate is relevant to future discussions on the census.

The authors are with the Department of Epidemiology and Public Health, The Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne, England.

Requests for reprints should be sent to Raj Bhopal, MD, Department of Epidemiology and Public Health, The Medical School, University of Newcastle upon Tyne, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom (e-mail: r.s.bhopal@ncl.ac.uk).

This paper was accepted February 24, 1998. Editor's Note. See related comment by Fullilove (p 1297) in this issue.
papers gave only a Black–White contrast. Words for the latter group included White, Caucasian, Anglo, and nationalities of persons (e.g., US, British). Jones et al. found these terms for White populations in the American Journal of Epidemiology. White, predominantly White, Anglo, Caucasian. Adhieh and Hahn made similar observations on research in the American Journal of Public Health.

Journal editors, the guardians of language, might be expected to take the lead in clarifying terms. The International Committee of Medical Journal Editors published its first, purfatory statement on terminology for research on race and ethnicity in 1997: "The definition and relevance of race and ethnicity are ambiguous. Authors should be particularly careful about using these categories." In most research on race, ethnicity, and health published in English (which includes most such work in the United States and the United Kingdom), the comparison group has been the majority (White) population. In this paper we consider the complex issue in the context of naming the population against which the health of racial and ethnic minority groups is compared. This context prompted our analysis, although the ramifications are clearly much wider. (Whether White populations are the most appropriate ones for comparison in ethnicity and health research is beyond the scope of this paper.)

The Purposes of Race, Ethnicity, and Health Research

The biological concept of race, whereby human populations were divided into subspecies, was dominant from the early 19th century until its decline with the collapse of Nazi racism at the end of the Second World War. In retrospect, the biological concept of race was ill-defined, poorly understood, and invalid, and the science based on it needed sharper scientific criticism.

The concept of ethnicity is slowly replacing the concept of race. While race and ethnicity are separate, overlapping concepts, they are often used synonymously. For example, in their examination of systemic lupus erythematosus by race, Hopkinson et al. used census data on populations by ethnic group, and recent textbooks use both race and ethnicity, either in the title or in the text, without distinguishing the ideas. The points in this article apply to both race and ethnicity.

The overriding perceived benefit of race, ethnicity, and health research is a better understanding of the causes of disease, particularly the relative contributions of genetic and environmental factors. (Differences postulated to be due to genetic factors should, however, be demonstrated to be so by genetic studies.) Such research is also part of surveillance that helps develop a scientific basis for policy, as in the United Kingdom and the United States.

Studies have shown multiple differences between racial and ethnic groups in the pattern of disease and culture, which researchers and policymakers have emphasized as important to effective health service delivery. Studies have also drawn attention to communication barriers and racist attitudes. Ethnicity and health research has, therefore, a scientific purpose of elucidating the causes of diseases and the interplay between cultural factors and health, and a practical purpose of ensuring that services and policies are appropriate. While the terminology and concepts used ought to be purpose driven, in practice this does not happen, particularly with the word White. Despite criticism of the historical and current concepts and methods underpinning race, ethnicity, and health research, such work is increasing.

The Role of the Comparison Population

Ethnicity and health research, or at least the principles derived from such research, must bridge geographic boundaries and time periods to be scientifically useful. The scientific paradigm, strong in epidemiology, is to seek understanding through comparison. The comparison population eases, but does not solve, the problem of interpreting disease variation between racial and ethnic groups in different places and times. For example, it is difficult to compare coronary heart disease rates in "South Asian" populations across the world because of differences in disease diagnosis and recording. Observations that in many countries rates have been higher among "South Asians" than among the local comparison population have made some interpretation possible. Comparison populations also help researchers interpret changes in disease trends and assess whether the health of minority groups deviates from expectations. Finally, in the context of access to and quality of health care, the comparison population permits assessment of inequities by ethnic or racial group.

Terms

Table 1 summarizes the qualities of most of the terms for nonminority populations in race, ethnicity, and health research in the United States and the United Kingdom. The dictionary-derived meaning shown in the table comes mainly from the Shorter Oxford English Dictionary.

The term White has long served social, political, and everyday life and is embedded in scientific language. The US Constitution institutionalized racial categories by allocating to slaves (Black or mulatto), for the purpose of political representation, the value of three fifths of a free person and by giving no such value to "Indians not taxed." The categories of free person, slave, and Indians not taxed were rapidly replaced by the preexisting color-based social classification of White, Black, and Red. The categories White and Black remain key descriptors of American society and are the basis of administrative statistics. The term Black has been debated and changed in everyday and official language, but the term White has not. The use of the term White in the British 1991 census question on ethnicity has legitimized its use in British epidemiology.

The US Office of Management and Budget's Directive 15 defines a White person as a person having origins in any of the original peoples of Europe, North Africa, or the Middle East. Until recently, persons from India were considered White in the US census, and Middle Eastern people still are. In Britain, Middle Eastern and North African people would not be considered White, and Asian Indians have never been considered so.

The term White includes persons of Scottish, New Zealander, Greek, Spanish, English, Canadian, Welsh, Irish, and—in the United States—Iranian and Moroccan descent and has little value in gauging ethnicity or race. It encourages the division of society by skin color, reinforcing racial stereotyping, and hides a remarkable heterogeneity of cultures. In a study carried out in London, of 39 White patients, 7 were Greek or Turkish Cypriot, 5 were Irish, and 9 were non-British European origin.

The term Caucasian categorizes populations on the ill-defined basis of a common origin in the distant past in the Caucasus region of Central Europe. Most populations originating in India, Pakistan, and Bangladesh, for example, are Caucasian. Except as an erroneous euphemism for referring to persons of European descent, the word has little value in race, ethnicity, and health research. Freedman's plea that this term be rejected by science should be accepted.

The words Occidental and Western have similar meanings and are geographic concepts (meaning belonging to the West). The heterogeneity of such populations is also
TABLE 1—Analysis of Terms Currently in Use to Describe Nonminority Populations

<table>
<thead>
<tr>
<th>Term</th>
<th>Dictionary-Derived Meaning</th>
<th>Strengths</th>
<th>Weaknesses</th>
<th>Comments and Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Applied to those races (chiefly European or of European extraction) characterized by light</td>
<td>Used in Censuses Socially recognized and historically lasting concept</td>
<td>Used to describe heterogeneous populations Unrelated to ethnicity</td>
<td>Misnomer in practice refers to people of</td>
</tr>
<tr>
<td></td>
<td>complexion</td>
<td>Antithesis of the term Black</td>
<td>Geographic links are historical rather than contemporary</td>
<td>European origin with pale complexions</td>
</tr>
<tr>
<td>Caucasian</td>
<td>Indo-European; Blumenbach’s (1800) term for the White race of mankind, which he derived</td>
<td>Some relation to genetic composition</td>
<td>Used to describe heterogeneous populations No contemporary geographic</td>
<td>Means originating in the Caucasus region and</td>
</tr>
<tr>
<td></td>
<td>from the Caucasus</td>
<td>Defines populations by geographic origin in the distant past</td>
<td>link Unrelated to ethnicity</td>
<td>refers to Indo-Europeans</td>
</tr>
<tr>
<td>Occidental</td>
<td>Native or inhabitant of the Occident (West)</td>
<td>Geographically based</td>
<td>Used to describe heterogeneous populations Unrelated to ethnicity</td>
<td>Means belonging to the West (occident is</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>where the sun sets)</td>
</tr>
<tr>
<td>Western</td>
<td>Of or pertaining to the Western or European countries or races, as distinguished from the</td>
<td>Refers to a culture and place</td>
<td>Not geographically specific Used to describe heterogeneous populations</td>
<td>Abandon</td>
</tr>
<tr>
<td></td>
<td>Eastern or Oriental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>European</td>
<td>A native of Europe</td>
<td>Signifies geographic origin Purports to describe a culture (though some</td>
<td>Used to describe heterogeneous populations Ancestral origins may be</td>
<td>Comparable in breadth to terms such as</td>
</tr>
<tr>
<td></td>
<td></td>
<td>would dispute its validity)</td>
<td>difficult to ascertain</td>
<td>Chinese, South Asian</td>
</tr>
<tr>
<td>Europid</td>
<td>Not defined but clearly connotes origins in Europe</td>
<td>Clear geographic status New term, no past associations</td>
<td>Used to describe heterogeneous populations Ancestral origins may be</td>
<td>Unfamiliar term</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>difficult to ascertain</td>
<td>Comparable in breadth to terms such as</td>
</tr>
<tr>
<td>General population</td>
<td>Not defined, but epidemiological meaning is everyone in population being studied</td>
<td>Makes no assumptions about racial/ethnic origin Truly a whole population</td>
<td>Inaccurate unless it is a truly representative population</td>
<td>Chinese, South Asian</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Useful for international studies comparing large areas</td>
</tr>
<tr>
<td>Indigenous</td>
<td>Native or belonging naturally to a place Pertaining to natives, aborigines</td>
<td>Links to land and birthplace</td>
<td>Imprecise Conflicts concepts of place of birth, place of residence, and</td>
<td>Some in the nonminority groups are</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ancestry</td>
<td>nonindigenous; some in the minority groups</td>
</tr>
<tr>
<td>Native</td>
<td>One born in a place One belonging to a non-European and imperfectly civilized or savage</td>
<td>Links to land and birthplace</td>
<td>Historical connotations of being non-European Conflicts concepts of place</td>
<td>Similar to, and used synonymous with,</td>
</tr>
<tr>
<td></td>
<td>race</td>
<td></td>
<td>of birth, place of residence, and ancestry</td>
<td>indigenous treatises</td>
</tr>
<tr>
<td>Majority</td>
<td>Greater number or part</td>
<td>Does not presume ethnicity</td>
<td>Extremely broad and imprecise</td>
<td>Avoid if possible</td>
</tr>
<tr>
<td>Non-Asian, non-Chinese</td>
<td>Not defined but implies those not belonging to the group under study</td>
<td>Logically correct</td>
<td>Extremely broad and imprecise</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Preferred</td>
</tr>
<tr>
<td>Reference, control,</td>
<td>The standard against which a population being studied can be compared In science, a</td>
<td>Neutral terms Recognize purpose of the nonminority group in the research</td>
<td>Nature of the reference or control population is not self-evident</td>
<td></td>
</tr>
<tr>
<td>comparison</td>
<td>standard of comparison used to check inferences deduced from an experiment, by application of the method of difference To place together so as to note the similarities and differences of</td>
<td>Forces writer to describe population and clarify terminology of study or review</td>
<td>Could be misunderstood to mean closer matching than is actually carried out</td>
<td></td>
</tr>
</tbody>
</table>

European has a geographic meaning and a more general one. In international research comparing, say, World Health Organization health statistics from different regions, European has a specific meaning. In practice, the term refers to people of European ancestral origin (e.g., studies in America may refer to persons of European stock) rather than nationality or residence.

Freedman argues the merits of the unfamiliar term Europid. It denotes a geographic base, has no alternative meanings, is self-
evident, and is a single word. These facts do not resolve the heterogeneity issue: a study from Aberdeen and one from Athens could both use the words European and Europid to describe the local population. The only attraction of Europid is that it would require users to pause and question the label.

General population is a good description if a population is truly representative of that under study, as in the work of Ecob and Williams. In their research the minority population was compared with the whole community (including minorities), an unusual strategy that dilutes variations between groups. The terms indigenous and native have no exact definitions and have pejorative associations. It is difficult to judge when an individual or population becomes indigenous or native. Terms such as majority population and non-minority population (for example, non-Asian) are as broad as the terms above but permit fewer assumptions about the racial/ethnic composition of the population. These terms can be interpreted as referring to the White population.

The terms reference, control, and comparison populations are grounded in scientific method and lead to no assumptions about race and ethnicity and so mandate a description of the population by the authors, as recommended in recent guidelines to authors. They invite the question, What is the composition of the comparison population? They also focus thought on the need to compare like with like and hence on the purposes of research on ethnicity and race. We recommend the use of these terms, provided they are not taken to imply normality, and suggest that comparison population be the preferred term.

Conclusion: Toward Guiding Principles

In recent years researchers have followed administrative categories for race and ethnicity, even when these are acknowledged as having no scientific or anthropological validity (see references 14 and 15). This unsatisfactory state of affairs can be remedied only if scientists use the most specific term suitable to the purpose and context of the study and avoid pejorative words. Careful description of the characteristics of each population studied to make clear the basis of racial or ethnic classification (e.g., ancestry, geographic origin, birthplace, language, religion, migration history) is an essential starting point.

In comparative work including a group from the majority population, terms such as reference, control, or comparison population have advantages compared with terms such as White or European. They raise fewer expectations and prior assumptions and require the writer to provide detail on the populations studied, including their heterogeneity and origins. Editors and reviewers will find it easier to spot a lack of such information if these terms are used. Comparison population avoids the implication of a standard or norm associated with the terms control and reference; therefore we suggest that this be the preferred term for a White population used for comparison.

This approach will not solve the problem of how populations perceive themselves in society, nor is it a solution to the classification problems in the collection of statistics for social and administrative purposes. It should allow scientists to break free from a nomenclature developed for nonscientific purposes and to participate in conceptualizing the basis of the racial and ethnic groupings they use. Given that scientific use of a social category can be interpreted as an endorsement of its validity, avoidance of loose terminology in research might influence everyday language and counter the predominance of color as a means of grouping populations.

There are problems of poverty and excess disease in subgroups of the White population, which cannot be unearthed and tackled by using the label White. For example, the Irish-born and Scottish-born residents of England and Wales have recently been shown to have the highest standardized mortality ratios in England and Wales, higher than those of racial and ethnic minorities born in countries of the Caribbean and South Asia. Clearly, there are subgroups within the White community with special needs. The argument that the focus of race and ethnicity statistics should be on those with adverse health outcomes is a sound one. Clearly, it is not only ethnic groups of color who are in this position.

This paper widens debate on the issue of conceptualizing, categorizing, and naming racial and ethnic groups. This debate has been misperceived as an issue mainly for minority groups. This paper has focused on the terminology used for populations in comparative studies of the health of ethnic and racial minority groups, but there are many aspects that need work by other scholars, such as whether there is international understanding and agreement on the meaning of the term White and other commonly used labels; whether such agreement is achievable; the comparative health of population subgroups aggregated within the term White; developing a valid nomenclature for that population and its subgroups; and demonstrating that data on subgroups of the White population can be successful in improving the health status of worse-off subgroups. Work is also needed to define whether comparisons of the health of minority populations with that of the majority population are an appropriate foundation for ethnicity, race, and health research. We recognize the scope of these tasks. Their achievement is a long-term goal, but discussion needs to start now.

Acknowledgments

We thank Dr Wai-Ching Leung and anonymous referees for helpful comments on earlier drafts, and both Mrs Lorna Hutchinson and Mrs Carole Frazer for help with preparing the manuscript. This paper was completed while the first author was a visiting scholar at the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill. Assistance from the Department and the School, in both resources and ideas, is acknowledged.

References


**APHA REPRINT SERIES #3**

Another New Collection from the APHA Reprint Series

**Homelessness in America**

Edited by Daniel B. Herman and Ezra Susser

This volume contains some of the most influential research that has been conducted to date on the problem of homelessness and specific conditions that are associated with homelessness. The articles share the methodological rigor, and the relevance to public health policy and practice that have long been hallmarks of the American Journal of Public Health.

The collection is organized into three sections. The first section, "The Problem of Homelessness: Description, Scope, and Underlying Causes," describes the nature and scope of the problem of contemporary homelessness, and includes several editorials that address underlying causes and issue calls to action. The next section, "Individual-Level Risk Factors for Homelessness," includes several of the best designed studies conducted to date that illuminate the role played by individual-level factors in vulnerability to homelessness. The third section of the volume, "Health Conditions among Homeless People," addresses both a broad range of health conditions that disproportionately affect homeless people as well as issues of utilization and access to health care services.

Mail to:
American Public Health Association
Publications Sales
P.O. Box 753, Waldorf, MD 20604-0753
Tel: 301/893-1894; Fax: 301/843-0159

$30 ($25 for APHA Members), plus shipping & handling, 1998
150pp., Stock No. 808/HAOD98 TO Order: Call (301) 893-1894,
Fax:(301) 843-0159, Email: TASCO1@APHA.ORG,
Web Site: http://www.apha.org