

# Race and Ethnicity in Genetic Research

Pamela Sankar,<sup>1\*</sup> Mildred K. Cho,<sup>2</sup> and Joanna Mountain<sup>3</sup>

<sup>1</sup>Department of Medical Ethics; Center for Bioethics; Leonard Davis Institute for Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania

<sup>2</sup>Stanford Center for Biomedical Ethics, Stanford University, Stanford, California

<sup>3</sup>Department of Anthropological Sciences and Department of Genetics, Stanford University, Stanford, California

Received 24 April 2006; Accepted 14 September 2006

Use of race and ethnicity terms in genetic research continues to generate controversy. Despite differing opinions about their basis or relevance, there is some agreement that investigators using these terms should: explain why the terms or categories were used, define them carefully, and apply them consistently. An important question is whether these recommendations are reflected in practice. Here we addressed this question based on 330 randomly selected articles published between 2001 and 2004 that reported on genetic research and used one or more words from a defined list of race, ethnicity, or population terms. The recommendation that authors using race or ethnicity terms explain the

basis for assigning them to study populations was met infrequently (9.1%), and articles that used race and ethnicity as variables were no more likely than those that used them only to label a sample to provide these details. No article defined or discussed the concepts of race or ethnicity. With limited exceptions, current practice does not reflect repeated recommendations for using race or ethnicity terms in genetic research. This study provides a baseline against which to measure future trends. © 2007 Wiley-Liss, Inc.

**Key words:** race; ethnicity; genetic research; terminology

**How to cite this article:** Sankar P, Cho MK, Mountain J. 2007. Race and ethnicity in genetic research. *Am J Med Genet Part A* 143A:961–970.

## INTRODUCTION

Recently a number of authors have called for greater clarity and precision in the use of race and ethnicity terms in scientific publications [Sankar and Cho, 2002; Burchard et al., 2003; Cooper et al., 2003; Phimister, 2003; Anonymous, 2004; Anderson and Nickerson, 2005; Race Ethnicity and Genetics Working Group, 2005]. These calls maintain that race, ethnicity, and associated population specific terms, such as black and white, are inconsistently applied [Hahn, 1992; Osborne and Feit, 1992; Williams, 1994; Bhopal et al., 1997; Bhopal and Donaldson, 1998; Sankar and Cho, 2002], and compared to other common variables in research, especially vague [Lee, 2004]. Inadequate definitions of population terms hinder study replication and slow scientific progress [Pritchard and Rosenberg, 1999; Pereira et al., 2001; Hirschhorn et al., 2002; Kahn, 2003]. Some have even suggested this laxity could result in harm to patients by generating misconceptions about the relative benefits of treatments for different populations [Anonymous, 2001; Kahn, 2003].

Concerns about the consequences of leaving these particular terms ambiguous in genetic research extend beyond the standard arguments for precision in science. Using a term without defining it can

suggest that its meaning is self-evident and therefore requires no definition. Readers, when faced with a passage that uses a race or ethnicity term without explaining its meaning, will not draw a blank as they might if confronted with an obscure technical phrase. Instead, because race and to a lesser extent ethnicity have widely used popular meanings, readers are likely to supply their own definitions. Research has shown that these definitions are often informed by unexamined stereotypes [LaViest, 1996; Bhopal, 2002; Cooper et al., 2003], that are typically wrong [Terracciano et al., 2005], and rarely positive [Blair et al., 2004; Burgess et al., 2004; Maddox, 2004]. Research has also shown that the unconscious reliance on such stereotypes by health care practitioners may contribute to racial and ethnic disparities in medical treatment [Rathore et al., 2000; van Ryn, 2002; Geiger, 2003; Sankar, 2007].

Grant sponsor: U.S. Public Health Service; Grant number: R01 HG003191.

\*Correspondence to: Pamela Sankar, Ph.D., Center for Bioethics, University of Pennsylvania, 3401 Market Street, Suite 320, Philadelphia, PA 19130. E-mail: sankarp@mail.med.upenn.edu

DOI 10.1002/ajmg.a.31575

Terms that represent common, recognizable categories need little explanation; by not explaining what they mean by terms, authors reinforce the impression that race or ethnicity categories can be consistently distinguished and readily assigned. The impression that race and ethnicity categories are self-evident can be especially troublesome in the context of genetic research because it can “reinforce the simplistic view that race/ethnicity is the sole and/or direct cause” [Kaplan and Bennett, 2003] of a condition, especially those associated with health disparities [Braun, 2002; Kaplan and Bennett, 2003]. Failure to specify a causal mechanism when reporting findings that link health outcomes to racial or ethnic identity can exacerbate this problem [Anonymous, 2001; Kaplan and Bennett, 2003; Goldstein and Hirschhorn, 2004].

Agreement on the need for greater clarity in genetics research about the function of race and ethnicity terms has sometimes resulted in suggestions that seem indistinguishable from guidelines for good scientific writing. Authors are advised to explain: (1) how race or ethnicity terms fit into a study, stating their relevance to the study hypothesis or why researchers chose a particular named population as study subjects [Gambaro et al., 2000; Nebert and Menon, 2001; Sankar and Cho, 2002; Burchard et al., 2003; Cooper et al., 2003; Kaplan and Bennett, 2003; Phimister, 2003; Anonymous, 2004; Anderson and Nickerson, 2005; Race Ethnicity and Genetics Working Group, 2005]; (2) how race or ethnicity was assigned to subjects and from what categories assignments were made [Bennett, 1997; Sondik et al., 2000; Rebbeck and Sankar, 2005; Anonymous, 2005b]; and, (3) a study's limitations with reference to the populations to which results can be generalized [Osborne and Feit, 1992; Anonymous, 1993; Davis et al., 2001; Ioannadis et al., 2004].

These suggestions are potentially valuable, but warrant scrutiny. First, their basis is often unclear. They rely not on systematic analysis of how authors actually use race or ethnicity terms but on anecdotal evidence and on what commentators consider the logical corollaries of using these terms in genetics research. Second, they are very similar to suggestions made several times over the past 50 years [UNESCO, 1952; Littlefield et al., 1999], but never evaluated. Before embracing this round of suggestions it would be useful to know how past ones have fared. Systematic studies of how race and ethnicity terms have been used in genetic research however are scarce. In order to assess the nature and degree of the problem and the effectiveness of proposals to address it, such inquiries are necessary.

With this need in mind, we conducted a study to examine how race and ethnicity terms are used in publications on genetic research. Specifically, we sought to assess the extent that proposals to improve

clarity accurately reflect and address the current state of research and to provide a baseline against which to judge the effect of proposals yet to come.

We describe how articles in the biomedical research literature use race and ethnicity, and test hypotheses that articles that use race and ethnicity terms justify or explain the use. These findings represent the first stage in a three-part study that examines practices and trends in genetics research literature. The next two stages focus on patterns within more narrowly defined subcategories of genetic research and on a detailed analysis of the function and meaning of race and ethnicity terms within, rather than across, articles.

## MATERIALS AND METHODS

In this analysis we examine how scientific articles describing genetic research define and use race and ethnicity terms. We avoid defining race or ethnicity, and do not seek to propose such definitions. Furthermore, we do not distinguish conceptually between race and ethnicity. We use both terms together as a phrase, as in “race and ethnicity,” unless specifically discussing one of them. Although many have argued for replacing race with ethnicity [Harrison, 1995; Wilson, 2000; Kalow, 2001; Schwartz, 2001; Wood, 2001] and have provided distinct definitions of the two terms, proponents themselves have not agreed on these definitions. In practice, the definitions and the conceptual distinctions they are intended to embody are not upheld, and the two terms are often used as synonyms.

### Sample Selection

We obtained samples of journal articles indexed in MEDLINE by searching on the basis of race and ethnicity terms, population terms, and genetics terms (details below). We based the search on three sets of journals. The first two samples were drawn from journals with the highest impact factors (based on ISI journal citation reports [Thomson ISI, 2005b]) in the fields of (1) clinical research (which included a subset of cardiology journals) and (2) genetics. Many high impact factor genetics journals do not address the human population research of interest here. As a result, and in order to keep the impact factor ratings roughly equivalent in the clinical and genetics journals, the high impact genetics sample is based on only two journals while the high impact clinical sample draws on five (as well as the cardiology journals). Table I lists the journals selected by impact factor. The third sample, referred to here as the general journal sample, was drawn from MEDLINE-indexed journals, excluding all of those journals that were included in the high impact sample.

We relied on journal impact factors to choose our sample for three reasons. First, impact factor as the

TABLE I. Sampled Journals by Impact Factor

Sample sets (N)	Journals sampled (2003 impact factor [Thomson ISI, 2005b])
Clinical (131)	New England Journal of Medicine (34.833) Nature Medicine (30.555) JAMA (21.455) Lancet (18.316) Journal of Clinical Investigation (14.307) Circulation (11.164) Circulation Research (10.117) Journal of the American College of Cardiology (7.599)
Genetic (102)	Nature Genetics (26.494) American Journal of Human Genetics (11.602)
General (97)	MEDLINE indexed journals, excluding those listed above (N/A)

“measure of the frequency with which the ‘average article’ in a journal has been cited in a particular year or period” [Thomson ISI, 2005a], suggests that articles from high impact factor journals are more likely to be read by more people. Thus analyzing articles from these journals provides an account of the most common models available to researchers who use genetic research findings but who may themselves not have conducted or been experts in this kind of research. Second, as impact factor rankings also approximate a journal’s prestige [Thomson ISI, 2005a], and as prestige often translates into greater resources, journals with high impact factors are more likely to have paid editorial staff and the capacity to develop and apply strict editorial and peer review policies. This capacity suggests that the practices identified in these journals might be more likely to reflect preferred practices. Third, articles in high-impact journals are the most likely to be reported in the lay press, and thus clarity in the use and definitions of socially-charged terminology is important to assess in these journals.

We added a specific sub-sample from high impact factor cardiology journals to the clinical sample to assure inclusion of a sufficient number of in-depth, condition-specific studies in the clinical sample. For analysis purposes, the cardiology sample was combined with the clinical sample.

To collect study articles we conducted a four step sampling process, applied first to high impact journal citations. We began step one by searching MEDLINE in the selected high impact factor journals with the following limits: abstract, humans, English, and publication dates between 2001 and 2004. To restrict articles to those that reported research, we also set search criteria to exclude articles indexed to the following article types: clinical conference, comment, consensus development conference NIH, duplicate publication, editorial, letter, news, newspaper article, review, academic, review, multicase, and review literature. Full citations for articles that met these criteria were then downloaded.

Step two started with searching the title, abstract and keywords of downloaded citations for terms from the following three lists: (1) Race and ethnicity terms, including race, racial, ethnic, ethnicity, and a set of race and ethnicity terms based on pre-2003 MEDLINE MeSH terms. These terms included: black, white, Caucasian, European-American, Asian, Hispanic American, Mexican American, Native American, American Indian, Alaskan American, African American, Inuit, Gypsy (or Gypsies) Arab, and Jew; (2) Population terms, including population, family, and kindred; and, (3) Genetic terms, including genetic (MeSH term “genetic” returns the following list of terms: cytogenetics, genetic research, genetics behavioral, genetics medical, genetics microbial, genetics population, immunogenetics, molecular biology, pharmacogenetics, and radiation genetics), and pharmacogenomics. It is noteworthy that MEDLINE has revised its MeSH terminology for the list 1 terms concerning race and ethnicity [Sankar, 2003]. However, list 1 terms were accepted MeSH terms during most of the period covered by our sample and their use by MEDLINE suggests a sufficient overlap with popular usage to warrant adoption here.

To be eligible for random selection for the study, an article had to have in its title, abstract, or keywords one term from list 1 (race or ethnicity terms) *or* one term from list 2 (population terms), *and* one term from list 3 (genetics terms). Review of article titles, abstract, and keywords based on these criteria returned 2,151 articles from high impact clinical journals, 178 from cardiology journals, and 1,295 from genetics journals.

For step three, we randomly selected 120 articles from the clinical set, 120 from the genetic set, and 45 from the cardiology set.

Step four subjected each of these articles to detailed review to assure that the article: (1) reported on original research; (2) used human subjects or human tissue samples; and, (3) concerned human genetics (and not bacterial or viral DNA). If an article did not meet one of the three criteria, it was excluded from the sample. Articles were also evaluated to assure that common words, such as “white” or “race” were used to describe the race or ethnicity of subjects. Articles that used these terms for different purposes or as part of another word were eliminated. (e.g., “white wine” [Mukamal et al., 2003], or “Arabidopsis” [Housworth and Stahl, 2003].) After this review, 100 articles remained in the clinical set, 31 from cardiology, and 102 in the genetics set. The cardiology articles were merged into the clinical set at this point resulting in a clinical sample of 131 articles.

To create the general article sample, we followed the same four steps with one difference. We ran the same MeSH and MeSH-based term searches, and selected articles based on the same combinations: a

race or ethnicity term *and* a genetic term, *or* a population term *and* a genetic term. However, from this set we then eliminated all articles from journals we had included in the high impact sample. Searches on the remaining articles returned 1,575 selections from which we then randomly selected 120 articles. Step four reduced this sample for the general article set to 97. Added to the high impact clinical and high impact genetics articles, these additional 97 articles resulted in a total sample of 330 articles for analysis.

### Coding and Inter-Rater Reliability

Once each sample was finalized, full text PDFs were downloaded for each article, then converted to Rich Text Format documents and imported into Atlas.ti version 5 [Muhr, 2004], a qualitative analysis software package. Content codes were developed to reflect how the research population was discussed, and the structure and main components of the article. The initial set of codes was defined and then tested by a team of four researchers, including PS and MKC. These codes were subjected to multiple evaluations including several rounds of consensus coding [Jenkins et al., 2005] and discussion, and review by two different advisory groups that included science journal editors and clinical and research geneticists. When the codes were judged to adequately capture the relevant article features, a codebook was created containing coding rules, definitions, and examples.

Three coders were trained using articles that had been coded as part of codebook development. Coder reliability was assessed on each article and training continued on additional previously coded articles until trainee coding largely matched approved coding. To evaluate inter-rater-reliability, these coders then coded a sample of 72 new articles. Each article was coded by a pair of coders and assignments were made such that each pair coded approximately one-half of the articles. Coding was then compared between coders and an agreement/disagreement ratio was calculated. The inter-rater-reliability score on these 72 articles was 90.6% agreement. Subsequently, each article was coded by one coder. 10% of the remaining sample was subjected to the same inter-rater reliability test and agreement remained >90%.

### Coding

Analysis codes are divided into 3 categories: (1) basic article features; (2) reasons researchers gave for how and why they used particular populations; and (3) the role of race or ethnicity in the research design or account of the research.

**Basic features.** Each article was classified on the basis of 3 basic features providing essential

information about the study it reported. Each code was applied as a dichotomous variable. (1) *Hypothesis* was defined as the presence of a founding idea or assumption stated as the basis for investigation. Text identified for this code included formally stated hypotheses and more general research questions. In either case, the text had to state or imply that the idea provided the basis for the study; (2) *Limitations* were statements that explained the factors that restricted the generalizability of study findings. To be coded as limitations, statements had to be explicit and related to study design. Hypothesis and limitations are standard features of scientific research articles that impose a structure on argument and inference [Gambaro et al., 2000; Ioannadis et al., 2004]. Inclusion of a specific hypothesis is important because this is where readers might expect to find an explanation of how the choice to characterize study populations as members of particular racial or ethnic groups relates to the study premise or hypothesis. A limitations statement provides the opportunity to explain how widely findings can be applied to populations beyond the study sample that might be associated with the race and ethnicity terms used in the study; and, (3) We defined *sample origin* not in geographic terms, but as where and how researchers acquired the tissue samples, for example, if they were already in researchers' possession, collected at a clinic, or obtained from a tissue bank.

**Reason for using populations.** To study authors' explanations of why research was conducted using race or ethnicity terms, articles were classified based on three features that have been recommended by expert commentary, journals, and professional societies [American Association of Physical Anthropologists, 1996; Anonymous, 1996; American Anthropological Association, 1998; American Academy of Pediatrics: Committee on Pediatric Research, 2000; American Anthropological Association, 2000; Rivara and Finberg, 2001; Winker, 2004; Anonymous, 2005a.]: (1) *Why population* was used to label text that stated reasons for pursuing the research question by using a population so identified; (2) *Why this population* was used to code reasons given for studying the particular population(s) in question. Reasons were practical (e.g., because the sample was available) or theoretical (e.g., because the condition of interest was known to occur frequently in a particular group); and, (3) *Basis for assigning population term* was defined as the way in which the label for the sample population was determined or assigned; for example, self-assigned by subjects, taken from existing records, or assumed because of the region where the participants were recruited.

**The role of race or ethnicity.** To examine different ways that articles used race or ethnicity, we labeled text that referred to the following 7 codes: (1) *Reference population* was applied to text where race

or ethnicity terms were used to name populations that were discussed in the article but not studied as part of the reported research. These statements typically occurred in the introduction or discussion sections; (2) *Sample population* was applied to text where race or ethnicity terms were used to label the study population; (3) *Dependent* and; (4) *Independent* were applied respectively to text that characterized race or ethnicity as dependent or independent variables in the research being reported; (5) *DNA with label* indicated where authors had labeled DNA, alleles, chromosomes, or mutations with a race or ethnicity term, as in “For the three variants combined, the difference between African and white chromosomes is significant” [Thye et al., 2003]. The 1st and 2nd codes, *reference population* and *sample*, and the 5th, *DNA with label*, could co-occur with the 3rd and 4th, but the 3rd and 4th were mutually exclusive; (6) *Defines race or ethnicity*. If an article had text coded to any of the first 5 codes, coders assessed whether to apply the code, *defines race or ethnicity*. This code was applied to text that defined race or ethnicity as a genetic, social, or biological concept; and, (7) *No race term*. By default, all articles that did not receive one of these 6 codes were classified *No race term*. These articles were found to contain one of the population terms used to select the sample (e.g., population, family, kindred).

Some articles referred readers to previously published articles for details about research procedures or the sample population. Such statements, if relevant to study codes, were double-coded, first to the primary code to which the statement referred, such as sample origin or basis for assigning population label, and then to a second code, *previously described*. After coding was complete, coders reviewed articles that displayed passages coded to previously described, noted the references cited

as containing the relevant information, and then located and downloaded these articles. The statements in referenced articles were then used as the basis for assigning codes to the original article.

### Statistical Analyses

We subdivided the data in two ways: according to journal category (clinical, genetic, and general), and temporally (first 2 years vs. last 2 years). We then compared frequencies of individual codes across the resulting subsets of the data, assessing the significance of any differences via the chi-square statistic. Statistical tests were performed using STATA statistical software (Version 7.0 for MAC).

### RESULTS

We reviewed and coded 330 articles describing genetics research in the context of human groups and found that slightly more than half used a race or ethnicity term as a variable (independent variable + dependent variable) and the rest divided almost evenly into those that did not use such a term at all for the reported study (no term + reference population only) and those that used it only to name the study sample and not as a variable (sample population only) (Table II). We compared articles from the first 2 years and the second 2 years and found no significant difference ( $P=0.664$ ).

### Basic Article Features

Most articles did not have clearly identifiable hypotheses or limitations. Coders were able to identify a hypothesis (passages that described an idea or premise as the basis for the investigation) in 30% ( $n=99$ ) of articles. Many articles in the sample followed a style of inquiry that can best be described

TABLE II. Coding Frequencies by Sample Set, N (%)

Variables coded	Clinical sample set (N = 131)	Genetic sample set (N = 102)	General sample (set N = 97)	Total (N = 330)
Basic features				
Hypothesis	50 (38.2)	21 (20.6)	28 (28.9)	99 (30.0)
Limitations	46 (35.1)	12 (11.8)	17 (17.5)	75 (22.7)
Sample origin	93 (71.0)	58 (56.9)	55 (56.7)	206 (62.4)
Reason for using population				
Basis for assigning population label	13 (9.9)	9 (8.8)	8 (8.2)	30 (9.1)
Why this population	17 (13.0)	12 (11.8)	8 (8.2)	37 (11.2)
Why population	17 (13.0)	9 (8.8)	10 (10.3)	36 (10.9)
Ways of using race and ethnicity terms				
No term	22 (16.8)	14 (13.7)	15 (15.5)	51 (15.5)
Reference population only	18 (13.7)	5 (4.9)	10.3% (10)	10.0% (33)
Sample population only	26 (19.8)	30 (29.4)	20 (20.6)	76 (23.0)
Independent variable	61 (46.6)	46 (45.1)	47 (48.5)	154 (46.7)
Dependent variable	4 (3.1)	7 (6.9)	5 (5.2)	16 (4.8)
DNA with label	3 (2.3)	22 (21.6)	10 (10.3)	35 (10.6)
Defines race or ethnicity	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)

as deductive rather than inductive in that the justification for the research seemed to emerge from its findings rather than from the question it posed. Other articles simply reported interesting findings with few details about why the research had been conducted. Neither type of article was coded to *hypothesis*. The number of articles providing a limitations statement was similarly low, 22.7% ( $n = 75$ ). Statements noting the origin of study samples was reported in 62.4% of the articles ( $n = 206$ ).

### Reason for Using Populations

The majority of articles did not justify the use of a labeled population or explain the basis of a population label. Just over 10% ( $n = 36$ ) of articles explained why research had been conducted using a labeled population, while slightly more explained precisely why a particular population had been chosen for study. Explaining the basis for how a label had been assigned to a population was identified in only 9.1% of articles ( $n = 30$ ). This last finding is particularly important because this feature most closely approximates what some guidelines have outlined as basic procedures for describing the identity of populations used in genetic research and might be considered a basic, easily fulfilled requirement.

Those authors who did explain the basis for assigning a label were likely also to explain why the research question they were pursuing was best studied using a population identified by a race or ethnicity term ( $P < 0.0005$ ). This group was also more likely to explain why the particular population they used was chosen, although the relationship was non-significant ( $P = 0.110$ ). The weak relationship between “explaining the basis for assigning a label” and “explaining why a particular population was chosen” may reflect the fact that in some studies the choice of a particular sample was driven more by convenience (access to samples) than by theory.

### Ways of Using Race and Ethnicity

Approximately half, 51.5%, of the articles used race or ethnicity terms as either a dependent or independent variable ( $n = 170$ ). Another quarter either did not use race or ethnicity terms at all or used them only to refer to populations discussed in related research

rather than to label their own study populations ( $n = 84$ ). The remaining articles (23%,  $n = 76$ ) used a race or ethnicity term to label members of the study population, but did not report results using this term nor test any hypothesis related to it. Just over 10% of the articles used phrases coded as *DNA with label* such as “Puerto Rican mutations” [Mykytyn et al., 2003], “African and Asian haplotypes” [Hernstadt et al., 2002], or more general phrases such as an “allele’s ethnic identity” [Collins-Schramm et al., 2002]. Most of these passages (22 of 35) appear in genetics journals.

No article in this sample defined race or ethnicity. Based on preliminary research we had anticipated that few articles would provide such a definition, but deemed it important to track nonetheless. We had not anticipated that no article would offer a definition. However, this finding is consistent with the low frequency of articles coded to *basis for assigning population label* in the sense that if a researcher neglects explaining how a population label was assigned, he or she is unlikely to define or discuss related concepts of race and ethnicity.

### Basic Article Features Associated With use of Race and Ethnicity Terms

If authors who used race or ethnicity terms as variables in their studies followed published recommendations, they would be more likely than authors who did not to clearly state study hypotheses and study limitations, and to explain the origin of study samples. To test these relationships we created a new *use of race or ethnicity term* variable by combining no use of terms (*no term*) with use of terms only to refer to a population other than the study population (*reference population only*) and by combining use of term as a *dependent* variable and use of term as an *independent* variable. Use of term to describe study sample (*sample population only*) remained unchanged.

Table III reports comparisons of the resulting three-category variable and the four basic article features, and shows that articles that use race or ethnicity terms as variables are no more likely to state the study’s hypothesis or limitations than those articles that use race or ethnicity terms only to label a sample or not at all. Articles that used race or ethnicity as variables differed significantly from other

TABLE III. Presence of Coded Article Features by Ways of Using Race or Ethnicity Terms

Feature coded in articles	Use of race or ethnicity term in total Sample (N = 330)			$\chi^2$
	No race term or only for reference population N = 84 (%)	To describe study sample N = 76 (%)	Dependent or independent variable N = 170 (%)	
Sample origin	40 (47)	49 (64)	117 (69)	0.004
Hypothesis	26 (31)	19 (25)	54 (32)	0.551
Limitations	14 (17)	18 (24)	43 (25)	0.296

articles in the greater likelihood of explaining the origin of study samples ( $P = 0.004$ ).

A second set of hypotheses was relevant only to those articles that used a race or ethnicity term either to label a study population ( $n = 76$ ) or as a variable ( $n = 170$ ). Based on recommendations about such use, we tested the hypothesis that articles that used race or ethnicity as a variable would be more likely than articles that used the terms only to label a sample to explain why they had chosen to conduct their studies using racially or ethnically identified populations, to explain why they had chosen the particular populations to study that they did, and to explain the basis on which they had assigned race or ethnicity terms to samples or subjects. We also tested the hypothesis that they would be less likely to use phrases coded to *DNA with label*.

Results in Table IV show that the articles that used race and ethnicity as variables were no more likely than those that used the terms only to label a sample to explain why they chose to study particular populations or to provide the basis on which they assigned race and ethnicity labels to subjects. However, these articles were more likely to use phrases coded to *DNA with label* ( $P < 0.0001$ ).

#### Differences Across Journal Types (Clinical or Genetic) and by Journal Impact Factor

To explore possible differences associated with the type of journal in which an article appeared, we looked only at articles from high impact factor journals and compared those from clinical journals to those from genetics journals. These analyses were conducted only on those articles that used race or ethnicity terms and are reported in Table V.

Findings suggest that articles published in clinical journals differ from those published in genetic journals on several measures. Clinical journal articles are significantly more likely to report study limitations and sample origins, while articles in genetics journals are significantly more likely to use phrases coded to *DNA with label*. The two sets of journals do not differ on any of the measures associated with providing reasons for using a populations labeled with race or ethnicity terms.

To look for possible differences associated with journal impact factor, we combined articles from the

clinical and genetics high impact factor journals and contrasted them to the non-high impact journal articles (i.e., the general article sample). This analysis eliminated the differences that were apparent between the clinical and the genetics journal when only high impact factor journal articles were analyzed. The high impact and non-high impact journal sets do not differ significantly on any measure.

#### DISCUSSION

Our inquiry sought to contribute to the debate about the use of race and ethnicity in genetic research by examining actual research practices, as reflected in published studies.

The focus on research practices reflects the opinions expressed in commentaries concerning this issue that suggested that improved study design and more accurate reporting of study design and results might resolve at least some of the controversy concerning the use of race and ethnicity terms in genetic research.

While there are several excellent studies that examine how other fields use race and ethnicity [Hahn, 1992; Cooper, 1994; Senior and Bhopal, 1994; Hahn et al., 1996; Drevdahl et al., 2001; Thomas, 2002], this is the first study to our knowledge to quantify the use of race or ethnicity as variables in genetic research and to document that, although use is widespread, explanation of use is generally rare. In our sample, no article that used race or ethnicity as a general term or concept defined them and fewer than 10% explained the basis for population terms that they assigned to subjects, an easily fulfilled requirement providing valuable information. Only a minority included clear hypotheses and limitation statements and few explained how specifying the racial or ethnic identity of the study population helped to answer the research question at hand.

These findings lend credence to concerns that genetics researchers neither address what they mean by race and ethnicity terms nor explain the relevance of race or ethnicity to their research. Precisely why authors persist in practices that do not reflect existing recommendations for using race and ethnicity terms in genetics research is unclear. It may be that some disagree with the recommendations while others lack information about them. Researchers using data

TABLE IV. Articles That Use Race or Ethnicity Terms to Name a Study Sample Compared to Articles that Use Such Terms as Dependent or Independent Variables

Feature coded in articles	Use of race or ethnicity term to describe sample or as a variable		$\chi^2$
	To describe study sample N = 76 (%)	Dependent or independent variable N = 170 (%)	
Basis for assigning term	7 (9.2)	23 (13.5)	0.339
Why population	7 (9.2)	26 (15.2)	0.196
Why this population	12 (15.8)	23 (13.5)	0.639
DNA with label	1 (0.01)	33 (19.4)	0.000

TABLE V. Variation Between Types of Journals and Between Articles From High Impact Factor Versus Non-High Impact Factor Journals

Variables coded	Clinical N = 65 (%)	Genetic N = 53 (%)	$\chi^2$	High impact factor N = 118 (%)	Non-high impact factor N = 52 (%)	$\chi^2$
Feature coded in article						
Hypothesis	26 (40.0)	15 (28.3)	0.184	41 (37.4)	13 (25.0)	0.209
Limitations	26 (40.0)	5 (9.4)	0.000	31 (26.2)	12 (23.0)	0.659
Sample origin	54 (83.0)	31 (58.5)	0.003	85 (72.0)	32 (61.5)	0.173
DNA with label	2 (3.0)	21 (39.6)	0.000	23 (19.4)	10 (19.2)	0.968
Reason for using population						
Basis for assigning population label	11 (16.9)	6 (11.3)	0.389	17 (14.4)	6 (11.5)	0.614
Why population	13 (20.0)	5 (9.4)	0.112	18 (15.2)	8 (15.4)	0.983
Why this population	12 (18.4)	6 (11.3)	0.283	18 (15.2)	5 (9.6)	0.322

sets that were created and labeled before such issues became prominent might hesitate to change assigned labels. Alternatively, reasons might lie with journal editors who either do not endorse such policies or do not have the resources to enforce them. These are questions in need of study.

Whether using race and ethnicity terms without explaining them actually contributes to unintended social consequences such as stereotyping or reinforcing incorrect understandings of race as a natural human feature, as some have postulated [Anonymous, 2001; Cooper et al., 2003; Kahn, 2003], is unknown and is itself a complex research question also in need of study. Nonetheless, even if these practices are found not to contribute to unintended social consequences, they warrant attention. For example, as others have noted, failing to provide an accurate explanation of how labels are assigned to subjects impedes constructive use of study findings [Hirschhorn et al., 2002] and, shorthand phrases such as “Puerto Rican mutations” [Mykytyn et al., 2003] or an “allele’s ethnic identity” [Collins-Schramm et al., 2002] are likely to confuse rather than clarify putative relationships between genetic variation and geographical origins.

We note a few limitations of this study. First, our journal sample was not representative of all clinical or genetic journals. However, we selected those with the highest impact factors in order to examine articles that presumably would be the most influential in their respective fields, serve as models for the research community, and have the most exposure outside the research community. Second, our analysis did not distinguish between different types of genetic research, where race or ethnicity might be used in different ways. Creating mutually exclusive categories for this variable that coders could reliably apply was not possible because of the heterogeneity of research and the uncommon but not rare approach to some research problems that led authors to report on more than one type of research (more or less as the stages of a study) in a single article. Such a variable might have accounted for some of the differences we reported here and attributed to other factors.

Despite these limitations, this study provides important and useful information about how investigators are using race and ethnicity terms in conducting and reporting genetic research. Our results suggest that reason for concern about this use persists and that, with the partial exception of clinical journals with high impact factors, inadequate explanation of the meaning and purpose of race and ethnicity is widespread across journals. Authors have suggested that to increase the likelihood that genetic research that uses race and ethnicity terms can contribute fully and positively to future research and clinical practice, articles need to provide complete explanations about their reasons for using these terms and their methods of applying them. Authors, peer reviewers, and journal editors can play an important role in encouraging changes in practice. Our analysis codes could be used as a basis for assessing the reporting and use of race and ethnicity in the biomedical literature.

#### ACKNOWLEDGMENTS

This work was supported by a grant from the U.S. Public Health Service, R01 HG003191. The funding source played no role in study design, data collection, analysis, or interpretation, nor in the writing of the manuscript and the decision to submit this paper for publication.

#### REFERENCES

- American Academy of Pediatrics: Committee on Pediatric Research. 2000. Race/Ethnicity, gender, socioeconomic status—Research exploring their effects on child health: A subject review. *Pediatrics* 105:1349–1351.
- American Anthropological Association. 1998. Statement on “Race”. Washington, DC.
- American Anthropological Association. 2000. American Anthropological Association response to OMB Directive 15: Race and ethnic standards for federal statistics and administrative reporting.
- American Association of Physical Anthropologists. 1996. Statement on biological aspects of race. *Am J Phys Anthropol* 101:569–590.
- Anderson N, Nickerson K, 2005. Special issue: Genes, race, and psychology in the Genome era. *Am Psychol* 60:5–8.



- Anonymous. 1993. Use of race and ethnicity in public health surveillance: Summary of the CDC/ATSDR Workshop. *MMWR Morb Mortal Wkly Rep* 42:16.
- Anonymous. 1996. Ethnicity, race, and culture: Guidelines for research, audit, and publication. *BMJ* 312:1094.
- Anonymous. 2001. Genes, drugs and race. *Nat Genet* 29:239–240.
- Anonymous. 2004. The unexamined ‘Caucasian’. *Nat Genet* 36:541.
- Anonymous. 2005a. Author guidelines: Race, ethnicity, and nationality. *Paediatr Perinat Epidemiol* (<http://www.blackwellpublishing.com/journal.asp?ref=0269-5022>).
- Anonymous. 2005b. Instructions for Authors. *JAMA* 294:119–124.
- Bennett T. 1997. “Racial” and ethnic classification: Two steps forward and one step back? *Public Health Rep* 112:477–480.
- Bhopal R. 2002. Revisiting race/ethnicity as a variable in health research. *Am J Pub Health* 92:156–157.
- Bhopal R, Donaldson L. 1998. White, European, Western, Caucasian, or What? Inappropriate labeling in research on race, ethnicity, and health. *Am J Pub Health* 88:1303–1307.
- Bhopal R, Kohli H, Rankin J. 1997. Editors’ practice and views on terminology in ethnicity and health research. *Ethn Health* 2:223–227.
- Blair I, Judd C, Fallman J. 2004. The automaticity of race and Afrocentric facial features in social judgements. *J Pers Soc Psychol* 87:763–778.
- Braun L. 2002. Race, ethnicity, and health: Can genetics explain disparities? *Perspect Biol Med* 45:159–174.
- British Medical Journal. 1996. Style matters: Ethnicity, race, and culture: Guidelines for research, audit, and publication. *BMJ* 312:1094.
- Burchard EG, Ziv E, Coyle N, Gomez SL, Tang H, Karter AJ, Mountain JL, Perez-Stable EJ, Sheppard D, Risch N. 2003. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med* 348:1170–1175.
- Burgess D, Fu S, van Ryn M. 2004. Why do providers contribute to disparities and what can be done about it? *JGIM* 19:1154–1159.
- Collins-Schramm H, Phillips C, Operario D, Lee J, Weber J, Hanson R, Knowler W, Cooper R, Li H, Seldin M. 2002. Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *Am J Hum Genet* 70:737–750.
- Cooper R. 1994. A case study in the use of race and ethnicity in public health surveillance. *Public Health Rep* 109:46–52.
- Cooper RS, Kaufman JS, Ward R. 2003. Race and genomics. *N Engl J Med* 348:1166–1170.
- Davis M, Bruckman D, Cabana M, Clark S, Dombkowski K, Kemper A, Rushton J, Freed G. 2001. Constructive use of race and ethnicity variables. *Arch Pediatr Adolesc Med* 155:973.
- Drevdahl D, Taylor J, Phillips D. 2001. Race and ethnicity as variables in Nursing Research. *Nurs Res* 50:305–313.
- Gambaro G, Anglani F, D’Angelo A. 2000. Association studies of genetic polymorphisms and complex disease. *Lancet* 355:308–311.
- Geiger H. 2003. Racial and ethnic disparities in diagnosis and treatment: A review of the evidence and a consideration of causes. In: Smedley B, Stith A, Nelson A, editors. *Unequal treatment: Confronting racial and ethnic disparities in health care*. edition. Washington, DC: The National Academies Press. p 417–454.
- Goldstein D, Hirschhorn J. 2004. In genetic control of disease, does ‘race’ matter? *Nat Genet* 36:1243–1244.
- Hahn RA. 1992. The state of federal health statistics on racial and ethnic groups. *JAMA* 267:268–271.
- Hahn R, Truman B, Barker N. 1996. Identifying ancestry: The reliability of ancestral identifications in the United States by self, proxy, interviewer, and funeral director. *Epidemiology* 7:75–80.
- Harrison FV. 1995. The persistent power of “race” in the cultural and political economy of racism. *Annu Rev Anthropol* 24:47–74.
- Herrnstadt C, Elson J, Fahy E, Preston G, Turnbull D, Anderson C, Ghosh S, Olefsky J, Beal M, Davis R, Howell N. 2002. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* 70:1152–1171.
- Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. 2002. A comprehensive review of genetic association studies. *Genet Med* 4:45–61.
- Housworth E, Stahl F. 2003. Crossover interference in humans. *Am J Hum Genet* 73:188–197.
- Ioannadis J, Ntzani E, Trikalinos T. 2004. ‘Racial’ differences in genetic effects for complex diseases. *Nat Genet* 36:1312–1318.
- Jenkins G, Merz J, Sankar P. 2005. A qualitative study of women’s views on medical confidentiality. *J Med Ethics* 31:499–504.
- Kahn J. 2003. Getting the numbers right: Statistical mischief and racial profiling in heart failure research. *Perspect Biol Med* 46:473–483.
- Kalow W. 2001. Pharmacogenetics, pharmacogenomics, and pharmacobiology. *Clin Pharmacol* 70:1–4.
- Kaplan J, Bennett T. 2003. Use of race and ethnicity in biomedical publications. *JAMA* 289:2709–2716.
- LaViest T. 1996. Why we should continue to study race, but do a better job: An essay on race, racism and health. *Ethn Dis* 6:21–29.
- Lee SS-J. 2004. “Incidental findings” of race in pharmacogenomics and the infrastructure for finding differences in biomedical research. In: Singer E, Antonucci TC, editors. *Proceedings of the Conference on Health Disparities*. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. p 81–94.
- Littlefield A, Lieberman L, Reynolds LT. 1999. The debate over race: Thirty years and two centuries later; Part Two: Thirty years after the debate over race: The passing of the great consensus. In: Montagu A, editor. *Race and IQ*. New York: Oxford University Press. p 70–84.
- Maddox K. 2004. Perspectives on racial phenotypicity bias. *Pers Soc Psychol Rev* 8:383–401.
- Muhr T. 2004. User’s manual for ATLAS.ti. Berlin: Scientific Software Development GmbH.
- Mukamal K, Conigrave K, Mittleman M, Camargo C Jr, Stampfer M, Willett W, Rimm E. 2003. Roles of drinking pattern and type of alcohol consumed in Coronary Heart Disease in men. *N Engl J Med* 348:109–118.
- Mykityn K, Nishimura D, Searby C, Beck G, Bugge K, Haines H, Cornier A, Cox G, Fulton A, Carmi R, Iannaccone A, Jacobson S, Weleber R, Wright A, Riise R, Hennekam R, Luleci G, Berker-Karazum S, Biesecker L, Stone E, Sheffield V. 2003. Evaluation of complex inheritance involving the most common Bardet-Biedl syndrome locus (BBS1). *Am J Hum Genet* 72:429–437.
- Nebert DW, Menon A. 2001. Pharmacogenomics, ethnicity, and susceptibility genes. *Pharmacogenomics J* 1:19–22.
- Osborne NG, Feit MD. 1992. The use of race in medical research. *JAMA* 267:275–279.
- Pereira A, Mota G, Bensenor I, Lotufo P, Krieger J. 2001. Effect of race, genetic population structure, and genetic models in two locus association studies: Clustering of functional rennin-angiotensin system gene variants in hypertension association studies. *Braz J Med Biol Res* 34:1421–1428.
- Phimister EG. 2003. Medicine and the racial divide. *N Engl J Med* 348:1081–1082.
- Pritchard J, Rosenberg N. 1999. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 65:220–228.
- Race Ethnicity and Genetics Working Group. 2005. The use of racial, ethnic, and ancestral categories in human genetics research. *Am J Hum Genet* 77:519–532.
- Rathore S, Lenert L, Weinfurt K, Tinoco A, Taleghani C, Harless W, Schulman K. 2000. The effects of patient sex and race on

- medical students' ratings of quality of life. *Am J Med* 108:561–566.
- Rebbeck T, Sankar P. 2005. Ethnicity, ancestry, and race in molecular epidemiological research. *Cancer Epidemiol Biomarkers Prev* 14:2467–2471.
- Rivara F, Finberg L. 2001. Use of the terms *race* and *ethnicity*. *Arch Pediatr Adolesc Med* 155:119.
- Sankar P. 2003. MEDLINE definitions of race and ethnicity and their application to genetic research. *Nat Genet* 34:119.
- Sankar P. 2007. Hasty generalization and exaggerated certainties: Reporting genetic findings in health disparities research. *New Genetics and Society* (in press).
- Sankar P, Cho M. 2002. Toward a new vocabulary of human genetic variation. *Science* 298:1337–1338.
- Schwartz RS. 2001. Racial profiling in medical research. *N Engl J Med* 344:1392–1393.
- Senior P, Bhopal R. 1994. Ethnicity as a variable in epidemiological research. *BMJ* 309:327–330.
- Sondik E, Lucas J, Madans J, Smith S. 2000. Race/ethnicity and the 2000 census: Implications for public health. *Am J Public Health* 90:1709–1713.
- Terracciano A, Abdel-Khalek A, Adam N, Adamovova L, Ahn C, Ahn H, Alansari B, Alcalay L, Allik J, Angleitner A, Avia M, Ayerst L, Barbaranelli C, Beer A, Borg-Cunen M, Bratko D, Brunner-Sciara M, Budzinski L, Camart N, Dahourou D, De Fruyt F, de Lima M, del Pilar G, Diener E, Falzon R, Fernando K, Fickova E, Fischer R, Flores-Mendoza C, Ghayur M, Gulgoz S, Hagberg B, Halberstadt J, Halim M, Hrebickova M, Humrichouse J, Jensen H, Jovic D, Jonsson F, Khoury B, Klinkosz W, Knezevic G, Lauri M, Leibovich N, Martin T, Marusic I, Mastor K, Matsumoto D, McRorie M, Meshcheriakov B, Mortensen E, Munyae M, Nagy J, Nakazato K, Nansubuga F, Oishi S, Ojedokun A, Ostendorf F, Paulhus D, Pelevin S, Petot J, Podobnik N, Porrata J, Pramila V, Prentice G, Realo A, Reategui N, Rolland J, Rossier J, Ruch W, Rus V, Sanchez-Bernardos M, Schmidt V, Sciculna-Calleja S, Sekowski A, Shakespeare-Finch J, Shimonaka Y, Simonetti F, Sineshaw T, Siuta J, Smith P, Trapnell P, Trobst K, Wang L, Yik M, Zupancic A, McCrae R. 2005. National character does not reflect mean personality trait levels in 49 cultures. *Science* 310:96–100.
- Thomas N. 2002. Revisiting race/ethnicity as a variable in health research. *Am J Public Health* 92:156.
- Thomson ISI. 2005a. The Impact Factor.
- Thomson ISI. 2005b. Journal Citation Reports.
- Thye T, Burchard G, Nilius M, Muller-Myhsok B, Horstmann R. 2003. Genomewide linkage analysis identifies polymorphism in the human interferon-g receptor affecting *Helicobacter pylori* infection. *Am J Hum Genet* 72:448–453.
- UNESCO. 1952. What is race? Evidence from scientists. Paris: UNESCO.
- van Ryn M. 2002. Research on the provider contribution to race/ethnicity disparities in medical care. *Med Care* 40:1140–1151.
- Williams D. 1994. The concept of race in health services research: 1966–1990. *Health Serv Res* 29:261–274.
- Wilson W. 2000. Ethnicity and APS. *J Autoimmun* 15:153–155.
- Winker M. 2004. Measuring race and ethnicity: Why and how? *JAMA* 292:1612–1614.
- Wood A. 2001. Racial differences in response to drugs—pointers to genetic differences. *N Engl J Med* 344:1394–1396.