

## Methodologic Issues in Minority Health Research

---

# Race and Ethnicity in Research on Infant Mortality

Matthew R. Anderson, MD; Susan Moscou, FNP, MPH

**Background and Objectives:** *Race and ethnicity are variables frequently used in medical research. However, researchers employ race and ethnicity in different ways and with differing intent. This leads to confusion over the interpretation of racial or ethnic differences. This study sought to determine how race and ethnicity are used in research on infant mortality.* **Methods:** *We did a structured literature review of original research related to infant mortality published between January 1995 and June 1996 and indexed in the Core Contents section of MEDLINE.* **Results:** *The majority of articles (54%) mentioned race and ethnicity. US studies mentioned race or ethnicity more than non-US studies (80% versus 22%). Only one study defined the method used to determine the ethnicity of patients; no study defined race or the methodology used in determining patients' race. Researchers primarily used race and ethnicity to describe study populations. Some racial and ethnic identifiers may have been stigmatizing to the subjects studied. The second most common use of race or ethnicity was as a potential confounder. Only one article discussed racism as a contributing factor in infant mortality.* **Conclusions:** *There are several problems and ambiguities in the use of race and ethnicity in clinical research. Researchers who use racial or ethnic categories should do so for specified reasons and adopt clear definitions of the categories used.*

(Fam Med 1998;30(3):224-7.)

Race and ethnicity are variables frequently used in medical research. However, their use is fraught with contradiction and can lead to confusion for both researchers and those who interpret clinical literature.<sup>1</sup>

Researchers use race with differing intent because there is "no accepted scientific definition of race."<sup>2</sup> Many investigators consider race to have a biological basis and interpret racial differences as evidence of genetic variation among populations. This perspective persists, although recent evidence from human population genetics has shown that "classification into races has proved to be a futile exercise."<sup>3</sup> At the other end of the spectrum, many researchers use race but recognize that it is a socially constructed grouping.<sup>1,4</sup> Arguments have also been made for eliminating racial categories altogether.<sup>5</sup> These contrasting concepts about race lead to confusion about how to use these variables and how to interpret differences reported in clinical research.

By the mid-1900s, biologists proposed using ethnicity rather than race as a way to categorize variation in human populations.<sup>6</sup> Unfortunately, the concepts of race and ethnicity became confused in 1978 when the Office of Management and the Budget created "race/ethnicity" as a category in federal statistics.<sup>7</sup> Thus, while race and ethnicity should not be synonymous, they have often become so in practice.

Two recent studies illustrate these problems by showing how race is used in and conceptualized in epidemiological and health services research.<sup>8,9</sup> These authors noted that race was included without justification and that it was often undefined. Additionally, Jones et al observed that race-associated differences detected in research studies were frequently not explained.<sup>8</sup>

Our study describes how race and ethnicity were used in original research published in a broad spectrum of journals. We examined infant mortality because this is an area in which differences in mortality among racial and ethnic groups are regularly reported. Our purpose was not to assess why these differences exist but rather to evaluate how race and ethnicity are applied as research variables.

---

From the Department of Family Medicine, Albert Einstein College of Medicine (Dr Anderson); and Montefiore Medical Center (Ms Moscou), Bronx, NY.

**Methods**

We conducted a structured literature review to identify all original medical research on infant mortality published between January 1995 and June 1996. We searched all articles indexed under the term *infant mortality*. We restricted our search to articles published in English and in the Core Contents section of MEDLINE.

Each article was reviewed independently using a recording instrument that we developed. Discrepancies between the two reviewers were resolved by refining the definitions used in the instrument.

The instrument recorded the following information for each reviewed article: 1) where the research was done (US versus non-US), 2) whether race and ethnicity were mentioned, 3) how race and ethnicity were defined, determined, and used, 4) what categories other than race or ethnicity were used by authors to conceptualize population differences, and 5) if racism was mentioned. We also examined four risk factors known to affect infant mortality (smoking, gestational diabetes, hypertension, and the adequacy of prenatal care) to see if race and ethnicity were mentioned more or less often than other risk factors.

**Results**

The literature search found 80 articles on infant mortality published between January 1995 and June 1996. (List available from authors on request.) Race and ethnicity were mentioned in 54% of all the articles (80% of US studies and 22% of non-US studies). Race and ethnicity were mentioned more often than the other four risk factors for infant mortality (Table 1).

*How Do Researchers Define or Determine Race or Ethnicity?*

Only one article, written by a non-US researcher, defined ethnicity.<sup>10</sup> No article defined race. Researchers extracted racial and ethnic data from hospital charts, perinatal databases, and vital statistics records such as birth certificates. No article provided specific information about what methods were used to determine the race of individual patients.

*How Do Researchers Conceptualize Population Differences?*

Table 2 displays the categories used by researchers to characterize population differences. The four categories used were race, ethnicity, nationality, and geo-

Table 1

Frequency of References to Race/Ethnicity and Four Selected Risk Factors for Infant Mortality

Mentioned	All Articles (n=80)	US Articles (n=44)	Non-US Articles (n=36)
Race	43 (54%)	35 (80%)	8 (22%)
Prenatal care	23 (29%)	18 (41%)	5 (14%)
Hypertension	21 (26%)	13 (30%)	8 (11%)
Gestational diabetes	10 (13%)	6 (14%)	4 (11%)
Cigarette smoking	16 (20%)	9 (20%)	7 (17%)

graphic origin. The majority of articles used either racial categories alone (44%) or racial and ethnic categories combined (30%).

*Why Do Researchers Use Race and Ethnicity?*

Table 3 describes reasons researchers use race or ethnicity. Race and ethnicity were used to describe the study population in 53% of the articles. Race and ethnicity were considered as potentially confounding variables in 42% of the articles. When race and ethnicity were seen as confounding variables, researchers factored them out of the study in the following ways: data presented indicated that the racial and ethnic composition of the study groups (eg, placebo and active treatment groups) were equivalent; race and ethnicity were used in a case-control match; and race and ethnicity were factored out using a logistic regression method. Race and ethnicity were central to the main hypothesis in 11 of the 80 articles (14%).

Table 2

Categories Used to Characterize Population Differences in 43 Articles Discussing Race/Ethnicity

	# of Articles	% of Total
Race only	19	44
Ethnicity only	1	2
Race and ethnicity	13	30
Race, ethnicity, and nationality	7	16
Race and geography	2	5
Race, ethnicity, geography, and nationality	1	2

**DEFINITIONS:**

**Race:** white, black, Asian or Pacific Islander, American Indian or Alaskan Native (as per US vital statistics<sup>7</sup>)

**Ethnicity:** Hispanic is considered the only ethnic category for US studies.<sup>7</sup> In non-US research, we accepted authors' characterizations of populations as "ethnic" groups.

**Nationality:** national origin, including hyphenated nationalities such as Japanese-Americans or Mexican-Americans

**Geographical group:** groups such as Southeast Asians or Central Americans. Asians were considered a racial category for US studies.

### Was Racism Mentioned?

Only one author (US) regarded racism, specifically residential segregation, as a potential factor in infant mortality.<sup>11</sup>

### Discussion

Our study documents that race and ethnicity are commonly used as variables in medical research related to infant mortality. Research performed in the United States mentioned race and ethnicity 80% of the time. In contrast, prenatal care is mentioned only 41% of the time, and maternal smoking is mentioned only 20% of the time. Further, race and ethnicity were central to the main hypothesis in 14% of studies.

How should race and ethnicity be used as variables? The Centers for Disease Control (CDC) sponsored a workshop in 1993 on "The Use of Race and Ethnicity in Public Health Surveillance."<sup>12</sup> Recommendations from the workshop included that researchers establish definitions of race and ethnicity, that patients or study populations self-identify their race and ethnicity, that researchers clearly indicate their reason(s) for analyzing data on race and ethnicity, and that studies document the effect(s) of racism.

Despite these recommendations, we found that no researchers defined race, and only one defined ethnicity. Although the CDC recommends self-identification as the most reliable method for determination of race and ethnicity,<sup>12</sup> researchers did not address how they determined the race or ethnicity of their study populations. Without clear definitions and methods for determining race and ethnicity, it becomes impossible to judge if results are valid. Further, the lack of definitions allows investigators to mix conceptual categories that do not necessarily belong together. Of the 43 articles discussing race and ethnicity, 23 intermixed racial data with data on ethnic groups, nationalities, or geographic area of origin. The difficulty posed by such mixtures of terms is apparent if one tried to generate testable hypotheses making comparisons from such disparate categories as "Chinese" (a nationality), "Hispanic" (an ethnicity), "Central American" (a geographic location), and "white" (a race).<sup>13</sup>

What function do race and ethnicity play in these studies? Our review documented that race and ethnicity serve primarily to describe study populations and secondarily as a potentially confounding variable.

In the process of describing populations, we noticed that negative stereotyping about minority groups occurred. For example, two articles used descriptors that might be considered stigmatizing. One study described subjects as from a "predominantly inner-city, poor, and black maternal population."<sup>14</sup> Another paper reported its subjects as "largely Mexican-Hispanic immigrants and socioeconomically deprived African-

Table 3

### Role Played by Race/Ethnicity in 43 Articles Discussing Race and Ethnicity

	# of Articles	% of Total
Central to main hypothesis	11	26
Describes study population	23	53
Potentially confounding variable	18	42
Simple variable	7	16
No data collected but was discussed	4	9
Risk factor	4	9

#### DEFINITIONS:

Main hypothesis: race or ethnicity was central to the main hypothesis of the investigation.

Descriptor: race or ethnicity described the study population.

Potentially confounding variable: racial and ethnic data was collected but factored out of the study so as not to confound the main thesis the authors were investigating.

Simple variable: racial or ethnic data was collected, analyzed, and reported.

Discussed, no data: studies mentioned race/ethnicity in the discussion section but did not collect data.

Risk factor: race or ethnicity was one element in an algorithm for determining perinatal risk.

Americans."<sup>15</sup> After making these descriptions, neither article contained any further reference to race or ethnicity. Many terms could potentially be used in describing any community. Why was it necessary to use these characteristics—poor, inner city, socioeconomically deprived, Mexican-Hispanic, immigrant, and black?

No article presented white patients with such descriptors. One study associated white patients with better "nutrition, exercise, and social habits," although there was no data to support this statement in the paper.<sup>16</sup> Similarly, a study compared predominantly black and Hispanic pregnant women cared for by midwives with a group of predominantly white prenatal patients cared for by obstetricians. In that study, the midwives' patients were chemically screened for substance abuse while the private practice patients were simply asked about substance abuse. While no private patient admitted to substance abuse, 10.3% of the midwife patients were diagnosed as abusing substances. This biased approach to screening invalidates the author's conclusion that patients seen by midwives had pregnancy outcomes similar to those of the private patients "in spite of . . . the 10.3% substance abuse" among the midwife patients.<sup>17</sup>

David and Collins contend that "the concept of race in medicine and epidemiology is more a result of habit and a reflection of existing power relationships than a product of critical scientific thought."<sup>18</sup> This was illustrated in two studies. One study examined infants with a specific biochemical defect (glucose-6-phosphate dehydrogenase deficiency) who resided in a

specific area of rural Nigeria.<sup>19</sup> Although the researchers had a precisely defined study population, they nonetheless reported that “all infants were black.” The other study compared children born at D.C. General Hospital in Washington, DC, to those born before arrival at the hospital.<sup>20</sup> The race of these two groups of infants did not differ significantly. Nonetheless, the researchers felt it necessary to state in the abstract that “Infants in both groups were predominantly black.” These examples reinforce the potential of race and ethnicity to stigmatize without providing useful information.

How might researchers better approach issues of race and ethnicity? We contend that when including racial and ethnic data as variables in a study, researchers should do so for a clearly stated purpose and should define race and ethnicity. They should clearly state the methods used for determining race and ethnicity (eg, self-identification) and demonstrate that those methods are reliable and valid. Comparisons between disparate conceptual categories such as nationality and race should not be made unless researchers give a clear rationale. We should require that the finding of differences between racial and ethnic groups lead to specific, testable hypotheses. Finally, as suggested by the CDC, racism should be considered as a contributing factor to racial and ethnic differences when such differences are found.

*Acknowledgment:* Preliminary data from this study were presented at the 1996 convention of the American Public Health Association in New York City.

*Corresponding Author:* Address correspondence to Dr Anderson, Albert Einstein College of Medicine, 360 E. 193rd Street, Bronx, NY 10458. 718-933-2400. Fax: 718-367-8168. E-mail: andersonma@aol.com.

## REFERENCES

1. Rushton JP. Do r-K strategies underlie human race differences? A reply to Weizman et al. *Canadian Psychology* 1991;32:1.
2. Osborne NG, Feit MD. The use of race in medical research. *JAMA* 1992;267:275-9.
3. Cavalli-Sforza LL, Menozzi P, Piazza A. The history and geography of human genes. Princeton, NJ: Princeton University Press, 1994.
4. King JC. The biology of race. Berkeley, Calif: University of California Press, 1981.
5. Montagu A. “The concept of race.” In: Montagu A, ed. The concept of race. London: Collier Books, 1964.
6. Huxley JS, Haddon AC. We Europeans: a survey of “racial” problems. London: Jonathan Cape, 1935.
7. Office of Management and the Budget. Race and ethnic standards for federal statistics and administrative reporting. *Federal Register* 1978;43:19269-70.
8. Jones CP, LaVeist TA, Lillie-Blanton M. “Race” in the epidemiologic literature: an examination of the *American Journal of Epidemiology*, 1921–1990. *Am J Epidemiol* 1991;134:1079-84.
9. Williams DR. The concept of race in *Health Services Research*: 1966 to 1990. *Health Serv Res* 1994;29:261-74.
10. Dhawan S. Birth weights of infants of first-generation Asian women in Britain, compared with second-generation Asian women. *BMJ* 1995;311:86-8.
11. Polednak AP. Trends in US urban black infant mortality by degree of residential segregation. *Am J Public Health* 1996;86:723-6.
12. Centers for Disease Control and Prevention. Use of race and ethnicity in public health surveillance. Summary of the CDC/ATSDR Workshop. *MMWR* 1993;42(no. RR-10).
13. Singh GK, Yu SM. Infant mortality in the United States: trends, differentials, and projections, 1950 through 2010. *Am J Public Health* 1995;85:957-64.
14. Meadow W, Reimshisel T, Lantos J. Birth weight-specific mortality for extremely low birth weight infants vanishes by 4 days of life: epidemiology and ethics in the neonatal intensive care unit. *Pediatrics* 1996;97:636-43.
15. Asakura H, Myers S. More than one previous cesarean delivery: a 5-year experience with 435 patients. *Obstet Gynecol* 1995;85:924-9.
16. Bianco A, Stone J, Lynch L, et al. Pregnancy outcome at age 40 and older. *Obstet Gynecol* 1996;87:917-22.
17. Blanchette H. Comparison of obstetric outcome of a primary care access clinic staffed by certified nurse-midwives and a private practice group of obstetricians in the same community. *Am J Obstet Gynecol* 1995;172:1864-71.
18. David RJ, Collins JW. Bad outcomes in black babies: race or racism? *Ethn Dis* 1991;1:236-44.
19. Slusher TM, Vreman HJ, McLaren DW, Lewison LJ, Brown AK, Stevenson DK. Glucose-6-phosphate dehydrogenase deficiency and carboxyhemoglobin concentrations associated with bilirubin-related morbidity and death in Nigerian infants. *J Pediatr* 1995;126:102-8.
20. Beeram M, Solarin K, Young M, Abedin M. Morbidity and mortality of infants born before arrival at the hospital. *Clin Pediatr* 1995;3(6):313-6.