9   Use of Racial and Ethnic Identity in Medical Evaluations and Treatments[1]

Jay S. Kaufman and Richard S. Cooper

The "evidenced based medicine" movement is premised on the transformation of clinical practice from a loose collection of traditions and opaque judgments into a modern science. The defining notions of a science in this context are that it should be based on systematic observation and experiment and that it should be self-correcting in the sense that conventions and beliefs are open to scrutiny and revision. Whereas medicine was once comfortably thought of as a somewhat mysterious application of the healing arts, the last half of the twentieth century saw an energetic movement toward placing medical practice within the realm of rational assessment and quantitative evaluation. Under this paradigm, it is no longer enough to apply a diagnostic test or surgical technique because of habit, tradition, or opinion. The promise of evidence-based medicine is therefore one in which the tyranny of professional judgment is replaced by the transparent authority of the data. While this movement has not been without its critics and controversies, there is no doubt that scientific principles and quantitative techniques have revolutionized modern medicine in our lifetime and have exposed many outdated medical myths and irrational practices to the bright light of critical evaluation (Sox et al. 2007; Sackett et al. 2000).

Despite the successes of this ongoing revolution, many holdovers from the nineteenth-century world are still very much a part of contemporary medical thinking, and perhaps none has proven more difficult to exorcise than the tenacious habits of racialized medicine. Although race has been thoroughly discredited as a meaningful biologic subdivision of humanity (Collins 2004; Torres and Kittles 2007), it is still a recurring and common quantity in medical training and practice. Despite grandiose twenty-first-century pronouncements of the declining significance of race and the emergence of a postracial society, it is still unremarkable to hear clinical cases presented on the basis of age, race, and sex for conditions that have no presumed relationship to racial identity, even if it were a valid descriptor in some sociological, if not biological, sense (Finucane and Carrese 1990; Caldwell and Popenoe 1995; Garcia 2004). Moreover, the tradition of racialized medicine has been vigorously defended

by some (Satel 2002), although many of the specific claims and assertions in these defenses have yet to be evaluated systematically or quantitatively. It is the purpose of this chapter, therefore, to outline a strategy for considering an evidentiary basis for when it would be rational and appropriate to include race as a consideration in medical evaluations and treatments.

Far from waning in the age of molecular genetics, race has been resurgent in biomedical discourse, especially in relation to a torrent of new interest in human biological variation and its quantification (Risch et al. 2002). There is little doubt that novel methodologies have ushered in a new era in our understanding of human evolution and variability, but the translation of this knowledge into sound clinical practice is what has gotten bogged down in a combination of old habits, reinvigorated superstitions, and outrageous promises. There are now not only race-specific guidelines for choice and dosage of existing therapies (Williams et al. 2004), for example, but even the advent of a wholly race-specific therapy (Sankar and Kahn 2005), with the promise of more on the way (Kahn 2007). The strange case of BiDil is exactly the kind of situation that clarifies why a methodically reasoned strategy is needed to assess such claims. Briefly, BiDil is a drug that was approved by the U.S. Food and Drug Administration (FDA) in June 2005 for treatment of heart failure in self-identified African Americans, on the basis of a successful randomized trial that recruited only self-identified African Americans (Taylor et al. 2004). Thus no formal claim of a race-specific effect was considered, and yet the therapy is labeled for use in only a single racial group. Clearly a systematic and sober approach to medical policy and practice regarding the use of race is, by now, long overdue. In this chapter, we attempt to outline what such an approach might look like, in broad terms, by working through a few simple examples and explaining the underlying logic employed.

Medical Decision Making

Medical decision making involves judgments under conditions of uncertainty and is therefore a naturally Bayesian exercise. To briefly explain the Bayesian paradigm, it is useful to contrast this with the frequentist paradigm that motivates much of the statistical analysis of data in contemporary biomedical sciences. The frequentist paradigm defines random variation as arising from the behavior of many repeated trials; for example, to say that a coin has a probability of landing on heads that is equal to 0.5 is a reference to the long-run proportion of heads over many repeated, identical flips. Frequentism is a natural paradigm for considering the behavior of a trial, such as a coin flipped many times, or the aggregate behavior of many electrons as they jump between energy shells around an atom. A weakness of frequentism, however, is that it cannot assign a probability to a single unrepeated event, such as the probability that average world average temperature will rise by five degrees within the next decade.

In the Bayesian statistical philosophy, however, probabilities are not proportions of repeated trials, but rather, statements of belief. A Bayesian begins with a prior probability distribution, which reflects some existing knowledge or belief. Then, in light of new data observed, the prior distribution is updated to form a new distribution, which is referred to as the posterior distribution. In this way, the Bayesian paradigm involves constantly changing one's degree of belief in light of new observations. This is exactly what should happen in the practice of clinical medicine (Weinstein and Fineberg 1980).

To make a diagnosis, prognosis, or decision regarding a potential medical intervention, a physician begins with prior knowledge. For example, when making a diagnosis, the prior probability distribution might be the frequency of the hypothesized disease in the source population. When seeing a patient with high fever in Africa, a physician might therefore assign a high prior probability to malaria as the underlying condition. The same encounter in the United States would instead assign a low prior probability to malaria, however, because the disease is rare in that setting. Next, the physician makes observations, for example, in the form of diagnostic tests, taking of history, and physical examination of signs and symptoms. These observed data serve to shift the clinician's prior beliefs, either toward confirmation or rejection of the original hypothesis. For example, observation of massive splenomegaly might move the physician much closer to certainty about the diagnosis of malaria. The goal of the exercise is to move close enough to certainty on either end of the belief spectrum that one can make a medical decision about the best course of action (figure 9.1).

It is not simple to state how close to one of the limits of certainty one must be to take some specific action, as this is a function of costs, risks, and patient's values. But the algebra for updating prior probabilities with the results of new observations can be quite straightforward, especially in the case of binary observations such as a positive or negative result for a diagnostic test. In that case, the basic formula is simply that posterior odds (O_2) is equal to prior odds (O_1) multiplied by the likelihood ratio of the test result (LR+ or LR–). The prior and posterior odds are simply $p/(1 – p)$, where p is the probability of the event of interest. For example, if prevalence of malaria is 20 percent, then $O_1 = 0.20/0.80 = 1/4$. For the result of a positive diagnostic test, the likelihood ratio positive LR+ is equal to the ratio of two conditional probabilities: the probability of a positive test result among those who truly have disease divided by the probability of a positive test result among those who truly do not have disease; that is, $\Pr(T+|D+)/\Pr(T+|D–)$, where the notation $\Pr(A|B)$ refers to the conditional probability of event A given event B. The definition of the likelihood ratio negative LR– is the similar ratio for that test result: $\Pr(T–|D+)/\Pr(T–|D–)$. This is all the algebra one needs to ask the simple question, how much does a given observed test result move us toward some degree of certainty? For example, if prevalence of malaria is 20 percent and splenomegaly has a LR+ = 5, then the observation of this symptom has the effect

Figure 9.1

Posterior probability of the hypothesis. The posterior probability guides subsequent action. If the patient is nearly certain to have the disease, the clinician will initiate treatment. If the patient is nearly certain not to have the disease, the clinician will reevaluate at some future time. If the probability is somewhere in the middle of this range from 0 to 1, then the clinician must make further tests or observations until the posterior probability is pushed to one side or the other.

of changing one's rational assessment of the probability of malaria from $O_1 = 1/4$ to $O_1 \times LR+ = 1/4 \times 5 = 1.25$. Therefore the posterior (i.e., updated) probability is now $1.25/2.25 = 0.56$. In light of the new evidence, the possibility of malaria is more likely than not, but the important feature to note is the role of the prior probability. If the same diagnostic test result were obtained in a low-prevalence setting, say, with prior probability of only 5 percent, then the posterior probability by the same calculation would be only 21 percent.

An Example: Screening for Sickle-Cell Trait

We now consider a simple example to show how this quantitative evaluation of evidence plays out when considering race as a possible determinant of medical diagnosis, prognosis, or treatment. The example revolves around screening newborn infants for sickle-cell trait, a condition that many would consider a paradigmatically racial trait (Wailoo 2001). Newborn children face a battery of screening tests, and so it is important to ask whether there are readily observable characteristics that might rule the newborn in or out for a particular screening test. For example, we do not routinely screen men for breast cancer, even though breast cancer does occasionally occur in men. Given the much lower occurrence of sickle-cell trait in nonblack infants, it was often proposed that screening be race-specific (Aspinall, Dyson, and Anionwu 2003).

We can therefore apply the logic outlined previously to evaluate such a proposal more formally.

The prior probability of sickle-cell trait for a black infant is more than tenfold higher than the corresponding probability for a white infant. Various surveys have estimated the prevalence of sickle-cell trait in the United States to be about 250 per 100,000 population in self-identified whites, and in the range of 6,500 to 7,000 per 100,000 population in self-identified blacks. Hispanics fall between these two values, with estimates around 500 per 100,000 population in the western United States and 3,000 per 100,000 population in the eastern United States. Asian and Native American populations have prevalences lower than whites, with estimates in the range of 100 to 200 per 100,000 population (Agency for Health Care Policy and Research 1993, Table 2). This might, at first, seem to favor the idea that race provides important information about the likelihood of uncovering a child with sickle-cell trait using a screening test, but it is instructive to subject the relevant numbers to the algebra described in the previous section before jumping to conclusions.

There were over 4 million children born in the United States in 2004, of which 2,304,181 were classified as "white, non-Hispanic" and 576,105 were classified as "black, non-Hispanic" (National Center for Health Statistics 2005). Using the sickle-cell trait proportions given earlier, this leads to 5,760 and 38,887 white and black births with sickle-cell trait in 2004, respectively. The calculation for the value of the likelihood ratio for white race is therefore (5,760/44,648)/(2,298,421/2,835,638) = 0.16. The corresponding likelihood ratio for black race is (38,887/44,648)/(537,218/2,835,638) = 4.6. We can now use these numbers to update the predictive probability of a newborn having sickle-cell trait, once we are able to observe and take account of the infant's race. The prior probability, before any race information is considered, is 0.0155 (i.e., between one and two cases per one hundred live births), so the prior odds is 0.0155/0.9845 = 0.0157. Suppose that we observe that the birth certificate records a child's race as black, and we take account of this information to help decide if this child should be screened. The formula is $O_1 \times LR_B = O_2$, which, in this scenario, leads to 0.0157 × 4.6 = 0.0722. Converting back from odds to probability, we obtain the posterior probability as 0.0722/1.0722 = 0.0673 (figure 9.2). To summarize, we began with an estimated probability for a newborn to have sickle-cell trait of about 1.5 percent, but taking into account that the child's race is categorized as black, we have updated that probability to 6.7 percent. While this increases the probability by over fourfold, it remains a relatively small probability. It would be difficult to specify costs and benefits of the screening program in such a way that it would be worthwhile to screen children with a 6.7 percent risk, but not children with a 1.5 percent risk.

What is more often considered, however, is ruling out the white infants. Similar calculations lead to a posterior odds of 0.0157 × 0.16 = 0.0025, which, converted to a probability, equals 0.0025/1.0025 = 0.0025. Therefore, taking into account that the

child's race is categorized as white, we have updated the probability from 1.5 percent to 0.25 percent. While this decreases the probability by about sixfold, the prior and posterior probabilities are both consistently small. If, under some cost and benefit structure, it is not worthwhile to screen children with an outcome probability of 0.25 percent, then it is probably not worth screening those with an outcome probability of 1.5 percent, either.

The overall conclusion here is that the use of racial information moves the initial probability a considerable distance in relative terms but a miniscule distance in absolute terms. For medical or public health treatment or policy decisions, therefore, the large prevalence difference in the population actually translates to a rather negligible quantity of information for making decisions about what to do with an individual. It is not surprising, therefore, that this conclusion was reached by an evidence review at the Agency for Health Care Policy and Research (AHCPR) in 1993. An expert panel reviewed similar information available at the time on prevalence of sickle-cell trait in the U.S. population by race and ethnic origin, to advise state screening programs (many of which had implemented race-specific targeting in their sickle-cell screening programs). In light of the limited information conveyed by race and ethnicity in this setting, however, and the high cost associated with missing a true case, the panel rec-

observation that child is black
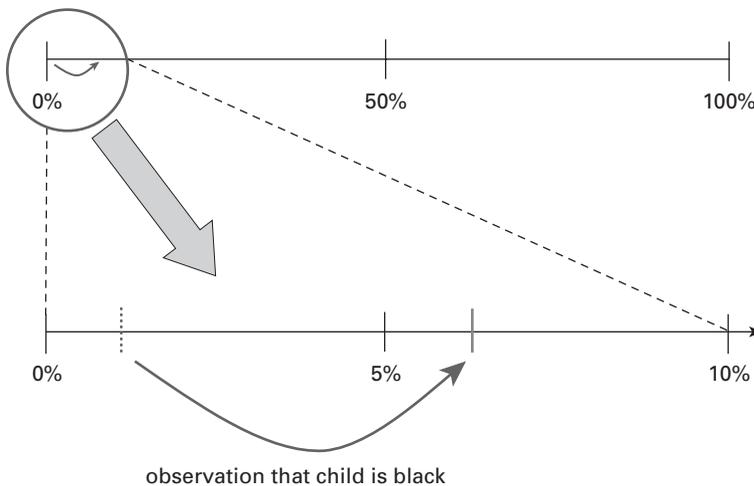
Figure 9.2
Updating probability for a black child. When trying to decide whether to screen a child after observing that his or her race is black, the probability shifts in light of race only on the far left-hand side of the probability range. This is magnified in the lower panel to show more clearly that the prior probability is 0.0155 and the posterior probability is 0.0673. While considering race increases the probability by over fourfold, it remains close to zero with or without accounting for this factor.

ommended universal screening of all newborns, regardless of racial or ethnic background, a recommendation that has now been implemented in all U.S. states (Agency for Health Care Policy and Research 1993).

Another Example: Choice of Initial Antihypertensive Therapy

Angiotensin-converting enzyme (ACE) inhibitors are a common first- or second-line antihypertensive therapy and have been widely and successfully used for this purpose for several decades. These drugs act on the renin-angiotensin system, which is the body's natural mechanism for raising blood pressure, by preventing the conversion of angiotensin I into angiotensin II. It is widely believed that ACE inhibition (along with other antihypertensive therapies that target the renin-angiotensin system) is less effective for black patients than for white patients, and this disparity in therapeutic efficacy is now described in virtually every hypertension textbook and set of clinical guidelines. For example, the "Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure" (JNC 7) represents the state of clinical knowledge on hypertension treatment in the United States (Chobanian et al. 2003). Although the document recommends first-line treatment with thiazide-type diuretics for all demographic groups, it also notes, under the heading "Other Special Situations," that "monotherapy with BBs, ACEIs, or ARBs lowers BP to a somewhat lesser degree in African Americans than Whites" (National Heart Lung and Blood Institute 2004, 39).

British Hypertension Society guidelines take a more extreme position with respect to this racial distinction, however, stating that blacks and those older than fifty-five years of age should have first-line treatment with calcium-channel blockers or thiazide-type diuretics, whereas whites and those under fifty-five years of age should be treated first with ACE inhibitors, angiotensin receptor blockers, or beta-blockers (Williams et al. 2004). The authors explain this differentiation by noting that hypertension can be classified into etiologic subtypes that are high renin or low renin, and that antihypertensive drugs that inhibit the renin-angiotensin system, such as ACE inhibitors, are less effective in low-renin hypertensives. They note that studies measuring plasma renin levels have reported that younger people and Caucasians have higher renin levels than do older people or blacks (which they define as people "of African descent"). This logic seems curious because the cited studies on plasma renin concentrations also indicate that such measures are not effective at predicting subsequent blood pressure response to treatment (Sagnella 2001).

Nonetheless, the guidelines are consistent with common medical wisdom, and similar recommendations can be found in innumerable textbooks and review articles. Gibbs, Beevers, and Lip (1999, 187), for example, note that "diuretics and calcium antagonists are suitable first-line agents in black hypertensives, whilst beta-blockers
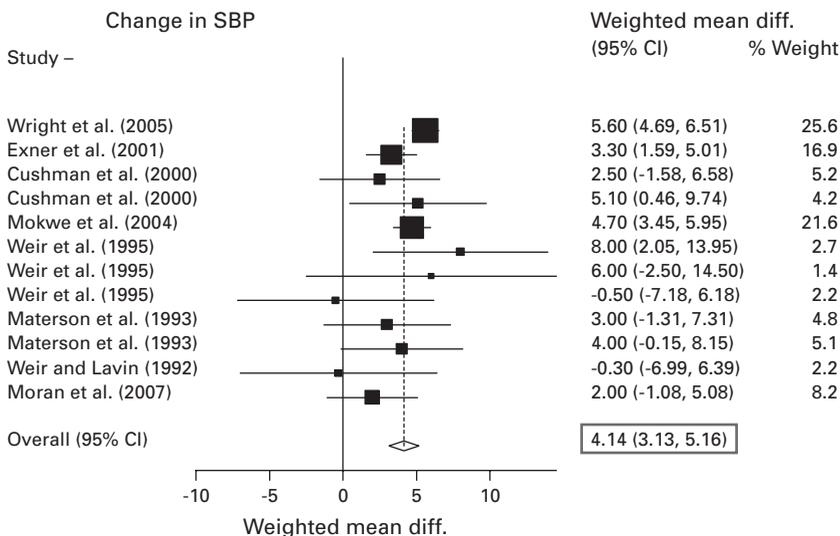
and the ACE inhibitors tend to be less effective at lowering blood pressure, due to the low renin state in these patients." Likewise, a comprehensive review of treatment recommendations by Douglas et al. (2003, 534) noted that ACE inhibitors had "some evidence of less blood pressure lowering efficacy in African American versus white patients." Moreover, these treatment recommendations are also reflected in drug package insert information. The ACE inhibitor trandolapril, for example, is marketed under the brand name "Mavik," and its package insert states that "recommended initial dosage of MAVIK for patients not receiving a diuretic is 1 mg once daily in non-black patients and 2 mg in black patients" (Abbott Laboratories 2003). The FDA announced, in 2003, that it had begun to ask drug manufacturers to report the race and ethnicity of all trial participants. In justifying this new requirement, the FDA noted several examples of race effects on drug treatment response, including the statement that "black people . . . have a poorer response than white people to beta blockers and angiotensin converting enzyme inhibitors" (Josefson 2003, 244).

We sought to evaluate the quantitative basis for differential treatment of hypertension by race, using published studies of blood pressure treatment responses to ACE inhibitors. Relevant studies were sought through PubMed MeSH (Medical Subject Heading) searches, and additional articles were found by reviewing the reference lists of the articles identified. In searching for studies using electronic databases of published reports, the following MeSH terms were used: "Angiotensin-Converting Enzyme Inhibitors," "African Continental Ancestry Group," and "European Continental Ancestry Group." In addition to these terms, we asked PubMed to identify only those articles that were in English and published in 1990 and onward. Two hundred twenty-two articles were initially identified. The search was last updated December 31, 2007. Studies were excluded if they were not randomized, if they involved children, or if they did not provide race-specific changes in systolic or diastolic blood pressure or the standard deviations of these changes. Studies were also excluded if multiple antihypertensive drugs were given at the same time or if information was not available for both blacks and whites. Of the 222 articles found, 19 met the inclusion criteria, and a total of 12 groups were available with all data necessary for conducting a meta-analysis of blood pressure treatment responses (Cushman et al. 2000; Exner et al. 2001; Materson et al. 1993; Mokwe et al. 2004; Moran et al. 2007; Weir and Lavin 1992; Weir et al. 1995; Wright et al. 2005). Some articles provided more than one group for analysis by subsetting the analysis on a study characteristic (e.g., urban vs. rural).

A random-effects meta-analysis of the existing studies was conducted using the STATA 9.0 statistical software package (Stata Coporation, College Station, Texas) and examining systolic and diastolic blood pressure responses. For simplicity, only results for systolic blood pressure are reported here. The presence of heterogeneity was assessed by examining the p value of Cochran's Q statistic, and there was no substantial heterogeneity detected in these study effects beyond what would be expected by

chance. The summary systolic blood pressure reduction difference between racial groups across these studies was estimated to be 4.14 mmHg, with a 95 percent confidence interval of 3.13 to 5.16 mmHg (figure 9.3). This value is largely consistent with the existing consensus in the literature, as described previously. Most large studies showed values that were very similar to this estimate, including the recently conducted ALLHAT trial (Wright et al. 2005).

Nonetheless, while the difference in population means was statistically significant in this summary over a large number of patients and studies, it was small compared to the within–racial group variability observed. Whites had a mean systolic decline in response to treatment of 10.7 mmHg (95% CI: 8.2, 13.3 mmHg), but the standard deviation of this value was 12.1 (95% CI: 8.9, 15.3 mmHg). Since these values have a roughly Gaussian (i.e., normal) probability distribution, this latter number indicates that about two-thirds of systolic blood pressure change values for whites obtained from a random sample would fall between –1.4 and 22.8 mmHg (i.e., ±1 standard deviation). For blacks, the mean systolic response was 6.8 mmHg (95% CI: 4.2, 9.4 mmHg), whereas the standard deviation for change in this group was estimated to be

Change in SBP	Weighted mean diff.	
Study –	(95% CI)	% Weight
Wright et al. (2005)	5.60 (4.69, 6.51)	25.6
Exner et al. (2001)	3.30 (1.59, 5.01)	16.9
Cushman et al. (2000)	2.50 (-1.58, 6.58)	5.2
Cushman et al. (2000)	5.10 (0.46, 9.74)	4.2
Mokwe et al. (2004)	4.70 (3.45, 5.95)	21.6
Weir et al. (1995)	8.00 (2.05, 13.95)	2.7
Weir et al. (1995)	6.00 (-2.50, 14.50)	1.4
Weir et al. (1995)	-0.50 (-7.18, 6.18)	2.2
Materson et al. (1993)	3.00 (-1.31, 7.31)	4.8
Materson et al. (1993)	4.00 (-0.15, 8.15)	5.1
Weir and Lavin (1992)	-0.30 (-6.99, 6.39)	2.2
Moran et al. (2007)	2.00 (-1.08, 5.08)	8.2
Overall (95% CI)	4.14 (3.13, 5.16)	

Weighted mean diff. scale: -10, -5, 0, 5, 10

Heterogeneity $x^2 = 16.9$ (df = 11); $p = 0.11$. Between-study variance $\tau^2 = 0.84$

Figure 9.3

White-black systolic blood pressure reduction. In a formal meta-analysis of all available studies that met the stated inclusion criteria, the summary white-black difference in response to angiotensin-converting enzyme (ACE) inhibitor monotherapy was 4.14 mmHg (shown as a diamond at the bottom of the figure). Studies were judged to be sufficiently homogeneous to pool in this way.

14.5 mmHg (95% CI: 10.7, 18.4 mmHg). Thus racial differences are quite small compared with the variation within each racial group (figure 9.4). This implies that knowing to which group a patient belongs tells a clinician little about the patient's individual response to treatment (Sehgal 2004, Nguyen et al 2009). As an analogy, consider that the mean height of men is significantly greater than the mean height of women, yet knowing the gender of a person helps very little in trying to estimate his or her individual height.

We can now translate these numbers into their potential impact on rational clinical decision making. If a clinician wanted, for example, a 10-mmHg systolic blood pressure decline in a white patient, the probability of achieving this decline would be obtained by integrating the Gaussian probability distribution function in figure 9.4 from $-\infty$ to 10, which yields a cumulative probability of about 52 percent. The same calculation for a black patient yields about 41 percent, giving a relative probability of success at this cutoff point roughly equal to 1.27. Repeating this calculation at every possible desired blood pressure decrement shows that the ratio reaches its maximum for a desired decline of about 20 to 25 mmHg, with a relative likelihood of success (favoring whites) of around 1.35. One could also compare densities instead of cumulative probabilities at each point, but the conclusion is approximately the same: a rela-
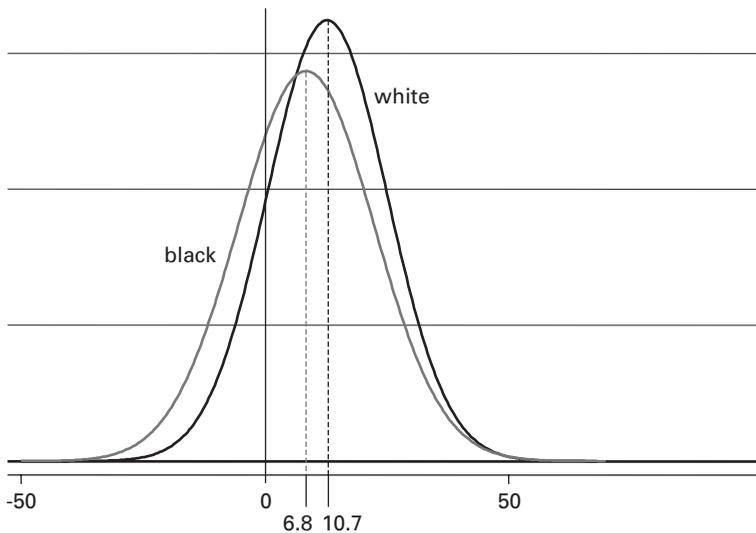
Figure 9.4

Systolic blood pressure reduction. Distributions of blood pressure response (in mmHg) for white and black adults receiving ACE inhibitor monotherapy are almost entirely overlapping. The means of the distributions may be "significantly" different, but knowing an individual's race tells one almost nothing about what that individual's blood pressure response will be.

tive benefit for whites in blood pressure reduction of as much as 30 percent to 40 percent. This relative benefit is also a likelihood ratio (i.e., the likelihood of successfully obtaining a desired systolic decline in whites relative to the likelihood in blacks). With a maximum value of less than 1.5, however, it is a very modest likelihood ratio in comparison with most clinical observations and diagnostic tests. A successful clinical indicator would separate the distributions of responders and nonresponders so that the clinician would have a high or low posterior probability that the patient would benefit from the treatment. It can be readily seen in figure 9.4, however, that at any desired blood pressure decline, there are many patients of each race group on either side of the cutoff point.

A substantial body of literature in medicine and epidemiology revolves around the large likelihood ratios necessary for individual inference in screening, diagnosis, or prognosis (Wald, Hackshaw, and Frost 1999; Pepe et al. 2004). For example, Pepe and colleagues note that if a potential marker (such as white race) misclassifies 10 percent of nonresponders as responders, the relative odds of successful response between whites and blacks is 3 (i.e., almost twice as large as the value in the case of ACE inhibition). This racial classifier can therefore identify only about 25 percent of the true responders; that is, 75 percent of the true responders are not going to be treated. Alternatively, if a marker of treatment susceptibility (such as white race) with a relative odds of 3 between whites and blacks can correctly classify 80 percent of the true responders, then it must mislabel almost 60 percent of the nonresponders as also being responders. Clearly this marker is not useful for individual-level prediction (Pepe et al. 2004).

Suppose a clinician had one hundred black and one hundred white patients and gave ACE inhibition only to the white patients. As shown previously, with respect to the values from our meta-analysis and a desired 10-mmHg decline in systolic blood pressure, the associated relative odds would be $(52 \times 59)/(41 \times 48) = 1.6$. So with respect to this 10-mmHg fall in systolic blood pressure, forty-eight of the white patients would fail to achieve the target blood pressure, and forty-one of the black patients would have achieved target blood pressure had they been given the drug, but instead, they were denied it. Therefore $(41 + 48)/200 = 89/200 = 45$ percent of the patients have been given the wrong treatment. From the point of view of any individual patient, this is not meaningfully better than being assigned by the flip of a coin.

Discussion

We have reviewed the basic notions of clinical decision making to assess how one would rationally decide when it might be appropriate to use racial identity as a clinical indicator. In a simple example involving screening for sickle-cell trait, we reviewed the logic that was used in recommending universal screening, which is now widely

adopted. In a second example, we examined a setting in which therapy is generally assigned on the basis of racial group in contemporary medical practice, and we showed that this current consensus does not appear to be well justified as a rational strategy. A reasonable critique of this second example, however, is that we showed only that the benefit associated with considering racial group is very small. If there is no cost to considering race, then any benefit, even a tiny one, could be considered to be better. After all, why not have a 52 percent chance of success, instead of a 41 percent chance? If an expensive or invasive test provided this benefit, it would be easy to show that the cost outweighs the benefit, but in the case of observed racial identity, the classification is made at no expense of time or materials. So why not make this observation and use that information?

The answer lies in the true costs of racial classification that arise from stereotyping. Racial groups are defined in terms of social and cultural affiliations and therefore are naturally imbued with the associations gleaned from everyday life in American society. This makes them a poor choice for serving as biomedical quantities because they naturally suggest to the clinician the patterns of traits and characteristics that are linked in the popular imagination to the group, even when these are inaccurate. For example, a 2006 study published in the *American Journal of Cardiology* found that black patients admitted with congestive heart failure at Grady Memorial Hospital in Atlanta, Georgia, were undertreated with beta-blockers, the current standard of care. The authors recommended more careful attention to current treatment guidelines (Ilksoy et al. 2006). A letter published subsequently in the journal by a clinician at George Washington University Medical Center argued against this recommendation, however, on the grounds that black patients would not take these medications, even if they were prescribed (Cheng 2006, 568): "African-American patients seem to be far less compliant with β blockers than diuretics due to the side effects of the former causing sexual impotence and weakness," he wrote. "A medicine is only good if the patient takes it as prescribed." The letter contains several citations on other points but fails to identify any previous literature that might support this claim of differential noncompliance for beta-blockers because blacks have greater concern about sexual function than other groups. It appears that the author found this point so self-evident that no supporting documentation was necessary.

Another cost associated with the use of racial classifications in medicine is when real differences that exist between groups are exaggerated in the calculations because they are treated like categorical distinctions. An anecdote of this type is provided by pediatrician Richard Garcia concerning a childhood friend who failed repeatedly to receive a correct diagnosis of cystic fibrosis until she was eight years old (Garcia 2004). Despite constant medical attention for pneumonia and other telltale symptoms, clinicians were unable to make the diagnosis of cystic fibrosis because she was black, and this observation was considered to be a counterindication. Garcia asserts that had she

been recognized as white, or even if she had been viewed as having no obvious racial identity of any kind, she would have been diagnosed immediately. Indeed, it was only when a radiologist, who had never seen her face-to-face, happened to see her chest X-ray that a correct diagnosis was made immediately. The problem in such circumstances is not that race can never be relevant for making diagnoses or other medical decisions—in fact, the prevalence of cystic fibrosis is indeed different in blacks and whites; rather, the potentially useful information conveyed by racial identity is simply much less than or much different than is commonly held, which puts many minority patients at a considerable disadvantage. The clinicians in Garcia's example, having learned that the prevalence of cystic fibrosis was lower in blacks, were acting as though the prevalence was therefore zero.

Moving beyond anecdotes like these, the overwhelming scientific evidence now suggests that these kinds of stereotypes and categorical elisions are commonplace in medical practice when it comes to clinicians trying to divine some information from patients' race (Smedley, Stith, and Nelson 2002). Various experimental trials have documented, under controlled conditions, the irrational and inappropriate use of racial identity in medical decision making. For example, Loring and Powell (1988) constructed dummy psychiatric case presentations that were intended to represent undifferentiated schizophrenia and labeled these with one of four race-sex combinations (black or white and male or female). The profiles were then assigned a diagnosis by psychiatrists, who returned questionnaires through the mail, with predictable results. Black patients, especially black men, were much more likely to be assigned the diagnosis of paranoid schizophrenia, indicating that clinicians perceived in these descriptions greater degrees of violence, suspiciousness, and dangerousness than for the identical white patients.

In a similar experimental design, Schulman and colleagues (1999) examined how often doctors recommended cardiac catheterization for hypothetical patients with chest pain. Physicians attending professional meetings were shown a videotaped interview with a "patient" (one of eight actors representing different combinations of race and gender) and were asked whether they would recommend catheterization. In addition to a significantly lower referral rate for black women, the authors reported significant differences by race and gender in the physicians' assessments of the patients' degree of hostility, negative affect, and socioeconomic status. Physicians also rated black patients as less likely to report symptoms, less likely to sue, and less likely to comply with treatment—all despite identical scripts read by the actors. The belief that minority patients are less likely to adhere to medication regimens or follow medical advice appears to be widely held, despite ambiguous evidence (Daniels, René, and Daniels 1994), and is used to justify undertreatment, as in the example of the heart failure treatment anecdote cited earlier. Beliefs that minority patients are less compliant or communicative leads to differential treatment that underserves these patients

by treating them closer to the presumed mean for their group. If the physician holds irrational beliefs about the group, this mean value may additionally be misspecified. But even in the best-case scenario, that the physician uses valid information about the "typical" patient from the relevant group, treating the patient more on the basis of this mean value, instead of the values obtained from the individual patient, will still create a systematic disadvantage for minority patients (Balsa and McGuire 2003).

We have argued thus far that there are rational ways to consider the amount of useful information contained in racial identifiers, and that race is used in medicine in many instances when it offers no substantial benefit and, furthermore, risks substantial harm. Given the fact that most biologic traits are largely, if not entirely, overlapping, as in the example of blood pressure response to ACE inhibition in the example given previously, we would hypothesize that justifiable uses of race in medical decision making will be rare. This leads to the natural question, however, about the scenarios in which it is likely to be useful. Although most biological variability is within group, rather than between group, for any subclassification of human beings by race or ethnicity, this is not true of social variability. While there are, for example, many tall and short individuals in every population, one could not say the same about social traits such as speaking Mandarin or eating tortillas. When disease arises from social behaviors that are highly patterned by ethnic group, racial or ethnic identity can be highly informative. Dietary examples are obvious, but social affiliations also matter for a variety of other conditions for which the use of race could be informative, even if controversial.

An example would be sexually transmitted diseases (STDs), where risk is a function not only of unsafe behaviors, but also of one's contact group (Aral 2000). Race is a major determinant of sexual contact patterns because of its overarching social significance. Indeed, sexual contact across racial lines only recently became legal in sixteen of the U.S. states after the 1967 Supreme Court decision of *Loving v. Virginia* (Hollinger 2003), and interracial marriages still accounted for only about 5 percent of registered unions in the 2000 U.S. census (U.S. Bureau of the Census 2000).[2] Given this situation of highly assortative sexual contact, it is not surprising that rates of many sexually transmitted infections vary dramatically by racial group. Prevalence of chlamydia among U.S. blacks is roughly eight times higher than the rate among whites, for example, whereas for gonorrhea, the prevalence is roughly nineteen times higher. Syphilis is seven times higher in blacks, which marks a decline from 1999, when the prevalence was twenty-nine times higher (Centers for Disease Control and Prevention 2009).

The origins of this dramatic prevalence disparity are rooted in a wide variety of long-standing social inequalities, from differential access to medical care to enormously unequal rates of incarceration. Once this pattern was established, however, it

was mostly endogamous partnership formation that operated to maintain such huge prevalence differences in STDs by race. Given such profound differences, reflecting the dominant role that race plays in U.S. society for social affiliations and residential patterns, it is not surprising that racial identity can play a legitimate (albeit uncomfortable) role in medical decision making around issues such as population screening for STDs.

For example, Stein et al. (2008) considered community-based testing for chlamydial infection in young adults under resource constraints. They found that using a small number of demographic characteristics as indications for screening, they could identify approximately 80 percent of infections, while testing less than half the population. However, the algorithm requires the use of racial identity because this variable is so highly predictive of infection. The reason that race turns out to be useful in this circumstance, but not in the earlier example of screening for sickle-cell trait, is also due, in part, to the higher absolute prevalence of chlamydia in the population when restricted by age and to the existence of other predictive covariates (e.g., number of sex partners in the last year), which led to even higher conditional probabilities. For ethical reasons, it may or may not be appropriate to use racial identity in this way. Nonetheless, with respect to the simple quantitative question of when race provides enough information to be of practical value, examples such as this one demonstrate that is more likely to occur when social recognition of race is a central etiologic factor.

Summary and Conclusion

It is well known that racial classification as practiced in Western society is only trivially correlated with genetic variability (Tishkoff and Kidd 2004). It is therefore unsurprising that biologic traits relevant for medical decision making tend to be largely overlapping across racial groups and that race is thus largely uninformative in diagnosis, prognosis, and treatment. Nonetheless, epidemiology and biostatistics provide a formal language for describing the quantity of information contained in a test or observation and a simple calculus for comparing a degree of belief before and after the introduction of this new information.

We have shown that even when the amount of information conveyed by race appears substantial, as in the case of sickle-cell trait, the use of this variable as a predictive tool may not be warranted. In a second example, involving antihypertensive treatment with ACE inhibition, we showed that although there is a reliably measurable difference in blood pressure response to the medication between racial groups, this difference is too small to be applied clinically to decisions involving an individual patient. Finally, after reviewing the limited benefits associated with use of race in medicine, we considered some of the costs in terms of stereotyping and the tendency

to misapply quantitative differences between groups as though they were categorical differences.

These considerations do not necessarily rule out the possibility of a thoughtful and appropriate use of race in medical care. As we suggested, when the disease of interest is etiologically related to social affiliations and environments, then race can be recognized as a real and consequential part of a patient's history and circumstances, with important implications for a wide range of experiences and exposures (Bhopal 2007). The existing track record for appropriate consideration of race in medicine is not very encouraging thus far, and one may doubt whether a social variable with such profound historical connotations can ever be used dispassionately, without invoking all the irrational debris of its sordid use as an instrument of social oppression. Nonetheless, the standard for twenty-first-century medicine is clear. Medical uses of racial classification should be viewed suspiciously, until they are justified with appropriately compelling theory and data. The traditional practice of using racially descriptive labels for patients, for example, should be reconsidered and, if found irrelevant, abandoned. Just as race-specific screening for sickle-cell trait was examined carefully and abandoned as ultimately unwarranted, so, too, should racialized guidelines for hypertension treatment, such as those of the British Hypertension Society, be abandoned, unless or until it can be shown that such recommendations provide a real benefit for patients and clinicians. The very first step as regards the use of such classifications would simply be to do no harm.

Notes

2. Population genetic analyses demonstrate that although the African American population is approximately 20% admixed with European ancestry, this gene flow occurred almost entirely from male Europeans during the institution of slavery (Parra et al. 1998).

References

Abbott Laboratories. 2003. Patient information for MAVIK. Reference 03-5264-R2. http://www.rxabbott.com/pdf/mavik.pdf.

Agency for Health Care Policy and Research. 1993. Sickle cell disease screening, diagnosis, management, and counseling in newborns and infants: Clinical Practice Guideline number 6. AHCPR Publication 93-0562. http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat6.section.17018.

Aral, S. O. 2000. Behavioral aspects of sexually transmitted diseases: Core groups and bridge populations. *Sexually Transmitted Diseases* 27:327–328.

Aspinall, P. J., S. M. Dyson, and E. N. Anionwu. 2003. The feasibility of using ethnicity as a primary tool for antenatal selective screening for sickle cell disorders: Pointers from the research evidence. *Social Science and Medicine* 56:285–297.

Balsa, A. I., and T. G. McGuire. 2003. Prejudice, clinical uncertainty and stereotyping as sources of health disparities. *Journal of Health Economics* 22:89–116.

Bhopal, R. 2007. *Ethnicity, race, and health in multicultural societies: Foundations for better epidemiology, public health, and health care*. Oxford: Oxford University Press.

Caldwell, S. H., and R. Popenoe. 1995. Perceptions and misperceptions of skin color. *Annals of Internal Medicine* 122:614–617.

Centers for Disease Control and Prevention. 2009. Trends in reportable sexually transmitted diseases in the United States, 2007. National Surveillance Data for Chlamydia, Gonorrhea, and Syphilis. http://www.cdc.gov/STD/stats07/trends.pdf.

Cheng, T. O. 2006. Beta blockers versus diuretics for congestive heart failure in African-American patients. *American Journal of Cardiology* 98:568.

Chobanian, A. V., et al. 2003. The seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: The JNC 7 report. *Journal of the American Medical Association* 289:2560–2572.

Collins, F. S. 2004. What we do and don't know about "race," "ethnicity," genetics and health at the dawn of the genome era. *Nature Genetics* 36:S13–S15.

Cushman, W. C., et al. 2000. Regional and racial differences in response to antihypertensive medication use in a randomized controlled trial of men with hypertension in the United States. Department of Veterans Affairs Cooperative Study Group on Antihypertensive Agents. *Archives of Internal Medicine* 160:825–831.

Daniels, D. E., A. A. René, and V. R. Daniels. 1994. Race: An explanation of patient compliance—fact or fiction? *Journal of the National Medical Association* 86:20–25.

Douglas, J. G., et al. 2003. Management of high blood pressure in African Americans: Consensus statement of the Hypertension in African Americans Working Group of the International Society on Hypertension in Blacks. *Archives of Internal Medicine* 163:525–541.

Exner, D. V., D. L. Dries, M. J. Domanski, and J. N. Cohn. 2001. Lesser response to angiotensin-converting-enzyme inhibitor therapy in black as compared with white patients with left ventricular dysfunction. *New England Journal of Medicine* 344:1351–1357.

Finucane, T. E., and J. A. Carrese. 1990. Racial bias in presentation of cases. *Journal of General Internal Medicine* 5:120–121.

Garcia, R. S. 2004. The misuse of race in medical diagnosis. *Pediatrics* 113:1394–1395.

Gibbs, C. R., D. G. Beevers, and G. Y. Lip. 1999. The management of hypertensive disease in black patients. *QJM* 92:187–192.

Hollinger, D. 2003. Amalgamation and hypodescent: The question of ethnoracial mixture in the history of the United States. *American Historical Review* 108:1363–1390.

Ilksoy, N., R. H. Moore, K. Easley, and T. A. Jacobson. 2006. Quality of care in African-American patients admitted for congestive heart failure at a university teaching hospital. *American Journal of Cardiology* 97:690–693.

Josefson, D. 2003. FDA calls for standardised racial and ethnic data. *British Medical Journal* 326:244.

Kahn, J. 2007. Race in a bottle. *Scientific American* 297:40–45.

Loring, M., and B. Powell. 1988. Gender, race, and DSM-III: A study of the objectivity of psychiatric diagnostic behavior. *Journal of Health and Social Behavior* 29:1–22.

Materson, B. J., D. J. Reda, W. C. Cushman, B. M. Massie, E. D. Freis, M. S. Kochar, R. J. Hamburger, et al. 1993. Single-drug therapy for hypertension in men. A comparison of six antihypertensive agents with placebo. The Department of Veterans Affairs Cooperative Study Group on Antihypertensive Agents. *New England Journal of Medicine* 328:914–921. Erratum, *New England Journal of Medicine* 330 (1994):1689.

Mokwe, E., S. E. Ohmit, S. A. Nasser, T. Shafi, E. Saunders, E. Crook, A. Dudley, and J. M. Flack. 2004. Determinants of blood pressure response to quinapril in black and white hypertensive patients: The Quinapril Titration Interval Management Evaluation trial. *Hypertension* 43:1202–1207.

Moran, A., J. A. Simon, S. Shiboski, T. G. Pickering, D. Waters, J. I. Rotter, C. Lyon, and D. Nickerson, et al. 2007. Differential effects of ramipril on ambulatory blood pressure in African Americans and Caucasians. *American Journal of Hypertension* 20:884–891.

National Center for Health Statistics. 2005. Demographics and health characteristics of births. *National Vital Statistics Reports*. http://www.infoplease.com/ipa/A0763852.html.

National Heart Lung and Blood Institute. 2004. The seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. NHLBI Report 04-5230. http://www.nhlbi.nih.gov/guidelines/hypertension/jnc7full.htm.

Nguyen, T. T., J. S. Kaufman, E.A. Whitsel, and R. S. Cooper. 2009. Racial differences in blood pressure response to calcium channel blocker monotherapy: A meta-analysis. *American Journal of Hypertension.* 22 (8): 911–917.

Parra, E. J., A. Marcini, J. Akey , J. Martinson, M. A. Batzer, R. Cooper, T. Forrester , D. B. Allison, R. Deka, R. E. Ferrell, M. D. Shriver.1998. Estimating African American admixture proportions by use of population-specific alleles. *American Journal of Human Genetics* 63:1839–1851.

Pepe, M. S., H. Janes, G. Longton, W. Leisenring, and P. Newcomb. 2004. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology* 159:882–890.

Risch, N., E. Burchard, E. Ziv, and H. Tang. 2002. Categorization of humans in biomedical research: Genes, race and disease. *Genome Biology*. http://genomebiology.com/content/pdf/gb-2002-3-7-comment2007.pdf.

Sackett, D. L., S. E. Straus, W. S. Richardson, W. Rosenberg, and R. B. Haynes. 2000. *Evidence-based medicine: How to practice and teach EBM*. Edinburgh: Churchill Livingstone.

Sagnella, G. A. 2001. Why is plasma renin activity lower in populations of African origin? *Journal of Human Hypertension* 15:17–25.

Sankar, P., and J. Kahn. 2005. BiDil: Race medicine or race marketing? *Health Affairs* W5:455–563.

Satel, S. 2002. I am a racially profiling doctor. *New York Times Magazine*, May 5.

Schulman, K. A., J. A. Berlin, W. Harless, J. F. Kerner, S. Sistrunk, B. J. Gersh, and R. Dubé, et al. 1999. The effect of race and sex on physicians' recommendations for cardiac catheterization. *New England Journal of Medicine* 340:618–626. Erratum, *New England Journal of Medicine* 340 (1999):1130.

Sehgal, A. R. 2004. Overlap between whites and blacks in response to antihypertensive drugs. *Hypertension* 43:566–572.

Smedley, B. D., A. Y. Stith, and A. R. Nelson, eds. 2002. Unequal treatment: Confronting racial and ethnic disparities in health care. Washington, DC: National Academy Press.

Sox, H. C., M. A. Blatt, M. C. Higgins, and K. I. Marton. 2007. *Medical decision making*. Philadelphia: American College of Physicians.

Stein, C. R., J. S. Kaufman, C. A. Ford, P. A. Leone, P. J. Feldblum, and W. C. Miller. 2008. Screening young adults for prevalent chlamydial infection in community settings. *Annals of Epidemiology* 18:560–571.

Taylor, A. L., S. Ziesche, C. Yancy, P. Carson, R. D'Agostino Jr., K. Ferdinand, M. Taylor, K. Adams, M. Sabolinski, M. Worcel, and J. N. Cohn; 2004. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. *New England Journal of Medicine* 351:2049–2057.

Tishkoff, S., and K. Kidd. 2004. Implications of biogeography of human populations for "race" and medicine. *Nature Genetics* 36:S21–S27.

Torres, J. B., and R. A. Kittles. 2007. The relationship between "race" and genetics in biomedical research. *Current Hypertension Reports* 9:196–201.

U.S. Bureau of the Census. 2000. U.S. census: PHC-T-19. Hispanic origin and race of coupled households: 2000. http://www.census.gov/population/cen2000/phc-t19/tab01.pdf.

Wailoo, K. 2001. *Dying in the City of the Blues: Sickle cell anemia and the politics of race and health*. Chapel Hill: University of North Carolina Press.

Wald, N. J., A. K. Hackshaw, and C. D. Frost. 1999. When can a risk factor be used as a worth-while screening test? *British Medical Journal* 319:1562–1565.

Weinstein, M. C., and H. V. Fineberg. 1980. *Clinical decision analysis*. Philadelphia: Saunders.

Weir, M. R., J. M. Gray, R. Paster, and E. Saunders. 1995. Differing mechanisms of action of angiotensin-converting enzyme inhibition in black and white hypertensive patients. The Tran-dolapril Multicenter Study Group. *Hypertension* 26:124–130.

Weir, M. R., and P. T. Lavin. 1992. Comparison of the efficacy and tolerability of Prinivil and Procardia XL in black and white hypertensive patients. *Clinical Therapeutics* 14:730–739.

Williams, B., N. R. Poulter, M. J. Brown, M. Davis, G. T. McInnes, J. F. Potter, P. S. Sever, S. McG. Thom. 2004. Guidelines for management of hypertension: Report of the Fourth Working Party of the British Hypertension Society, 2004-BHS IV. *Journal of Human Hypertension* 18:139–185.

Wright, J. T. Jr., J. K. Dunn, J. A. Cutler, B. R. Davis, W. C. Cushman, C. E. Ford, and L. J. Haywood, et al. 2005. *Journal of the American Medical Association* 293:1595–1608.