Race in Biological and Biomedical Research

Richard S. Cooper

Department of Public Health Sciences, Loyola University Medical School, Maywood, Illinois 60153

Correspondence: rcooper@lumc.edu

The concept of race has had a significant influence on research in human biology since the early 19th century. But race was given its meaning and social impact in the political sphere and subsequently intervened in science as a foreign concept, not grounded in the dominant empiricism of modern biology. The uses of race in science were therefore often disruptive and controversial; at times, science had to be retrofitted to accommodate race, and science in turn was often used to explain and justify race. This relationship was unstable in large part because race was about a phenomenon that could not be observed directly, being based on claims about the structure and function of genomic DNA. Over time, this relationship has been characterized by distinct phases, evolving from the inference of genetic effects based on the observed phenotype to the measurement of base-pair variation in DNA. Despite this fundamental advance in methodology, liabilities imposed by the dual political-empirical origins of race persist. On the one hand, an optimistic prediction can be made that just as geology made it possible to overturn the myth of the recent creation of the earth and evolution told us where the living world came from, molecular genetics will end the use of race in biology. At the same time, because race is fundamentally a political and not a scientific idea, it is possible that only a political intervention will relieve us of the burden of race.

A Klee painting named Angelus Novus shows an angel looking as though he is about to move away from something he is fixedly contemplating. His eyes are staring, his mouth is open, his wings are spread. His face is turned toward the past. Where we perceive a chain of events, he sees one single catastrophe, which keeps piling wreckage upon wreckage and hurls it in front of his feet. The angel would like to stay, awaken the dead, and make whole what has been smashed. But a storm is blowing from Paradise; it has got caught in his wings with such violence that he cannot close them. The storm irresistibly propels him into the future to which his back is turned, while the pile of debris in front of him grows skyward. This storm is what we call progress.

—Walter Benjamin

Theses on the Philosophy of History

Race as the Unwelcome Guest in the Disciplines of Science

We rarely appreciate the presence of history in our day-to-day experience. The quotidian is a mixture of the repetitive and the predictable, carried forward by habit and punctuated by random events that we regard as either good or bad fortune. But in a more reflective mood, we have to acknowledge the relentless force of history that holds us in its grasp and accept that it creates the possibilities we use to negotiate with the future. The imposition of racial categories on human populations has been one of the most enduring historical forces that sets limits on opportunity and thereby shapes our life tra-
jectory. As a projection of the underlying power relationships onto individuals, racial categories are used to structure social inequality. These power relationships are manifested both in the belief system that rank orders intrinsic human qualities according to group membership and the social institutions that enforce this hierarchy by restricting access to wealth, education, and other social goods. This daily reality is central to the history of all modern societies.

The racial structuring of society also has pervasive influence on biological research and the patterns of health and disease. Enormous effort has been expended to describe human demographic history through reference to an ever-changing array of constructs and categories, all of which include a hierarchical arrangement—either explicit or implicit. In the United States, most prominently, public health has embraced racial/ethnic categories as fundamental structural elements. Clinical medicine has similarly evoked racial categories to explain causation and outcomes across the entire spectrum of diseases. At the same time, race has met some of the strongest challenges to its legitimacy in biology and biomedicine. All of biology is grounded in the theory of descent from a common ancestor. The belief in racial categories was one of the most powerful liabilities of premodern biology and lent credence to the established view of divine creation. Indeed, it has recently been argued that the challenge to race brought by the abolitionist movement was a key factor behind Darwin’s transformative insight that the biological world—on the evolutionary time scale—is a single indivisible whole (Desmond and Moore 2009). Biomedicine still grapples with the implications of that insight for our species, yet substantial progress—uneven, tentative, and ultimately disappointing—has been made. In the current era, genomic science has opened new vistas onto previously unobserved dimensions of biology, and that proportion of the concept of race that has been attributed to genetics can finally be subjected to empirical scrutiny. Integrating this new knowledge into practice and focusing the technology on socially productive work, as always, remains our most difficult challenge.

The narrative of race therefore wanders the border territory between what we call science and what we recognize as history and politics. In the pregenomic era, there was no requirement—indeed, no opportunity—to validate the authority of race with molecular evidence; causal inferences were made on the basis of phenotype, in its broadest possible sense, from disease to accumulated material wealth to social graces. The primary purpose of the race concept was to serve as a shortcut, an organizing tool that allowed postenlightenment Europe to explain—and thereby justify—how imperialism had reshaped the world. Consequently, for both the social and biological sciences, race felt like the rude cousin whose claim on our affection was based on obligation, not choice. In every historical period, an incremental struggle has been waged to overcome the disruption that this unwelcome intruder has caused within empirical scientific disciplines.

In its origins, race was a “label of convenience” that biologists used interchangeably with the construct of “varieties” as they tried to create taxonomic categories below the level of the species (Cooper 1984). Writers from across the intellectual spectrum of literature and politics also felt free to make use of the idea. Thus, Baudelaire spoke of the “race of Abel” and the “race of Cain” when describing the polarization of 19th-century French society, and Marx characterized the English working class as a “race of peculiar commodity owners” (Baudelaire 1857; Marx 1957) (“Hence the sum of the means of subsistence necessary for the production of labour-power must include the means necessary for the labourer’s substitutes, i.e., his children, in order that this race of peculiar commodity-owners may perpetuate its appearance in the market.” [p. 172]). Malleability continues to be an essential quality of race, although it is now primarily used as a label for the temporary and often random aggregation of population subgroups, usually tied in some rough way to the perceived continent of origin (Kaufman and Cooper 1996). In its contemporary sense, biological race has now come to signify the inherited qualities of a population group hidden inside the DNA molecule.
The uses of race in biomedical research in the United States can be divided roughly into four historical periods. From the mid-1800s the dogma of racial inferiority meant that racial-genetic explanations were invoked as biological justification for discriminatory and genocidal policies (Montagu 1942; Cooper 1993a). The sea change that resulted from antifascist struggles of the 1930s and 1940s, together with emergent voices of African-American self-assertion, issued in a second phase and constrained the use of blatant biological determinist arguments. In the last half of the 20th century and extending into the present, the utility of race has been viewed from two distinct perspectives: as a descriptive category—necessary to document health inequalities—and as a causal explanation of ill health—through unspecified genetic influences (Cooper 1993b). Finally, in the contemporary era, the standard has shifted to require molecular evidence for causal effects that are to be ascribed to biological race; if this standard is not always met in practice, it can now at least be justified in theory.

The terms of the debate over “race and health” have therefore been sequentially recast. The first major challenge to claims of genetic causation was directed toward the “inference from phenotype” approach. Because most common diseases result from chronic exposures to noxious environmental stimuli, and because the pattern of exposure to the entire gamut of environmental exposures is highly structured by social class, geography, and other historical factors, disentangling an “intrinsic” property of race from the summed effect of poorly measured or unknown external risk factors is generally not feasible. In technical terms, because “race”—as a social category—is designated with relative precision, and lifetime exposure to the wide range of correlated risk exposures is characterized only approximately for individuals, causal analyses are undermined by residual confounding (for a more detailed discussion, see Kaufman et al. 1997; Cooper 2001). The “race” variable, and its imputed biological meaning, retains statistical significance that should most likely be distributed over other factors in the analytic model, or factors that remain unmeasured.

Typically, in studies comparing U.S. blacks and whites, income and education are the primary variables used to adjust for the influence of the social environment. On reflection, however, it is apparent that a whole universe of other factors, e.g., quality of education, neighborhood of residence, inherited wealth, interaction with the criminal justice system, etc., have a profound impact on life’s chances and physical and mental well-being. Although the logic of the argument for residual confounding has been generally accepted by epidemiologists, in practice it is more often ignored—perhaps in part because it identifies a problem without offering a solution. In fact, we will almost never have the capacity to summarize mathematically the impact of racial discrimination. Nietzsche recognized the philosophical implications of this problem, which is deeply rooted in many forms of belief: “The falseness of an opinion is not for us any objection to it... The question is, how far an opinion is life-furthering, life-preserving, ...and we are fundamentally inclined to maintain that the falsest opinions are indispensable to us, that without a recognition of logical fictions, without a comparison of reality with the purely imagined world of the absolute and immutable, without a constant counterfeiting of the world by means of numbers, man could not live” (Beyond Good and Evil, Chap. 1, epigram 3). Although usually interpreted as a criticism of Platonism and traditional religion, this characterization of useful fiction is equally apt when applied to belief in racial superiority, and it is precisely the “counterfeiting... by means of numbers” that concerns us here.

The ability to probe the genome has now changed the course of much of human biology, including such disciplines as evolution, epidemiology, demography, and clinical medicine. Genomics is also fundamentally altering how race is conceptualized and used in biomedical research. In the initial phases, the dominant constructs of the pregenomics era were simply carried forward. Large numbers of candidate gene studies, for example, purporting to show
a higher frequency of “risk alleles” in populations with increased disease rates—usually African Americans—were published from multiple disciplines. These studies were often interpreted as offering an explanation of public health inequalities. On technical grounds alone, this generation of candidate gene studies was ill conceived as a way of elucidating heritable influences and in fact yielded mountains of false-positive results; it was therefore easy to exploit this approach to satisfy the inherent biases of race. Fortunately, genome-wide association studies have proven to be more reliable, and for the first time we have credible insights into variation in genetic factors across a range of populations.

A vast amount of new information is being acquired with the tools now available, and although we are still in the early stages, for the first time the genetic architecture of many health-related traits is being described. As a consequence, we also now have the opportunity to accumulate data on the overarching question of the contribution of genes to racial patterns of disease. In parallel with this progress on disease-associated genetic variation, much deeper insight has been obtained into population genetics, quantifying relatedness and difference among racial/ethnic groupings.

All of the major causes of disability and death in the United States occur more frequently in blacks than in whites, with the exception of lung disease and suicide. Although all of these conditions have been analyzed from the basis of racial predisposition, none have been more central to the discourse about race and health in North America than hypertension. As a disorder that results from “dysregulation” of a basic physiologic system, hypertension has an ill-defined status, putting it somewhere between a “trait” and a “disease.” It cannot be biopsied, X-rayed, or localized to an organ or biochemical pathway. Conceptually, it reflects a free-floating quality of the whole organism, and its origin is therefore subject to a wide range of speculation, in many ways mirroring the essentialist attribute of race. This quality may underlie some of the unyielding temptation to ascribe higher risk of hypertension among blacks to a genetic cause (Cooper and Rotimi 1994; Cooper et al. 1999).

HIGH BLOOD PRESSURE AND AFRICANS IN THE NEW WORLD

By the 1930s, both clinical and population studies had shown that black Americans had higher average blood pressure than did persons of European descent (Cooper et al. 2000). The sequelae of hypertension—primarily stroke, heart disease, and renal failure—still account for most of the overall health disadvantage experienced by black adults in the United States. Among blacks older than age 55 in rural Georgia, mean systolic blood pressures approaching 180 mmHg were recorded in the 1950s. Because the researchers were unable to identify causal factors in the environment, the prevailing view was that “the frequency of hypertension in a given population...(was determined by)...the frequency in that population of a hypertensive gene” (White 1967, p. 183). Paul Dudley White, the preeminent American cardiologist of the first half of the 20th century, was particularly taken by the “high frequency with which hypertension has been found to exist among our American Negroes” (White 1967, p. 150). Once we are able to forgive the quaint manner of speech that White used to frame the question, it is undeniable that he formulated an important insight into this problem: “I used to ask my medical friends...about the situation in West Africa as to the prevalence of...hypertension in the jungle....” (White 1967, p. 150). But the evidence that White could rely on at that time was fragmentary, based on hospital cases from what was then Leopoldville in the Congo, and he concurred with the prevailing notion of “racial predisposition” (White 1967, p. 151). This perspective remains the dominant view until today.

As a medical student in the late 1960s, I too was deeply moved by the severity of hypertension among black patients I saw from the delta farmlands of eastern Arkansas. Devastating strokes in patients in their 40s and 50s were common. In the context of the 1960s, especially in Arkansas, there was good reason to question all
assumptions about race. The invocation of a mysterious essentialist quantity inherent in an individual with darker skin color—in the face of profound differences in life experience—seemed like nothing more than another way of formulating the “big lie” about the history of the south. At that time, of course, there was no way to test the validity of any of the conventional claims being made. In fact, through the 1990s, I still maintained that attempts to define the mechanism underlying black–white differences in blood pressure belonged in the realm of neo-science—the study of the unknown and the unknowable—rather than science (Cooper and Kaufman 1998).

As a result of interactions with colleagues from West Africa and the Caribbean, I was presented with the opportunity to test at least the “inference by phenotype” version of the black–white blood pressure story. Using neighborhood samples, my colleagues and I recruited 11,000 adults between the ages of 25 and 64 in urban and rural West Africa (Nigeria and Cameroon), the West Indies (Barbados, St. Lucia, and Jamaica) and the United States (metropolitan Chicago). Blood pressures in rural Africa were low and rose very little with age; in the West Indies, average blood pressures were the same as those found among whites in North America; the metropolitan Chicago sample (drawn from the community of Maywood), on the other hand, reproduced the U.S. data as a whole, with a 50% excess of hypertension, compared with whites (Fig. 1) (Cooper et al. 1997). As one might expect, the lifestyle factors known to increase blood pressure—namely, obesity, high intake of sodium, and low intake of potassium, fruits, and vegetables—paralleled the gradient in blood pressure. Subsequently, we compared these data to national surveys in the United States, Canada, and several European countries (Fig. 2) (Wolf-Maier et al. 2003). The “white” populations of Europe had substantially higher blood pressures than the non-U.S. Afro-origin populations, and in northern Europe levels were similar to or higher than that of the U.S. blacks study population (Cooper et al. 2005). Other studies have subsequently shown that immigrants to the United States from the Caribbean have lower blood pressures and lower stroke

![Figure 1. Prevalence of hypertension among six populations of West African origin. (From Cooper et al. 1997; reprinted, with permission, from the author.)](http://perspectivesinmedicine.cshlp.org/)

Cite this article as Cold Spring Harb Perspect Med 2013;3:a008573
rates than U.S.-born blacks. Clearly, therefore, the construct of “black exceptionalism” in terms of the risk of hypertension must be reconsidered as at most a “permissive susceptibility.” The true counterfactual state—persons of recent African descent who live the U.S. lifestyle without facing either racial or class disadvantage—does not exist, and although deprived of any affirmative evidence, the answer to the question of genetic predisposition remains unknown.

**GENOMICS AND STUDIES OF RACIAL FACTORS IN DISEASE**

Conclusions about racial predisposition based on the observation of difference in phenotype have therefore become untenable—although they continue to be routinely reported in some fields—and the complex task of sorting through the molecular evidence has begun (Cooper et al. 2003). As described elsewhere in this collection, one of the first constructs that required revision was the notion of “unit race”—that the various lists of population groups were cataloging similar segments of the world’s population. The evolutionary history of *Homo sapiens* was immediately apparent in any genetic studies that included African-origin groups. Our family tree is rooted in Africa, and the vast majority of accumulated genetic variation is present in contemporary African populations. Thus, Africans are no longer the “outlier” or the “other,” but in fact they are the source population and the reference point for all other continental groups.

As suggested above, the application of this new technology to the study of common disease followed a confused path through family studies, comparison of case and noncases at specific loci, and finally genome-wide association studies simultaneously comparing markers on all chromosomes. Statistical genetics had long since provided the basis for a strong assumption that quantitative traits were multigenic, that is, many mutations in many genes and their regulatory components were responsible for interindividual variation in phenotypes. Surprisingly, geneticists and epidemiologists alike opened this chapter of research with the opposite assumption that mutations of relatively large effects could be found. When integrated into a search for genetic factors underlying racial predisposition, the choice of assumption is obviously crucial. In modern societies, chronic diseases are the primary threat to health. Two-thirds of all deaths result from cardiovascular diseases and cancer alone. The pathophysiology of these conditions as a rule reflects dysregulation of metabolic pathways (e.g., lipid metabolism, blood pressure control, and DNA repair mechanisms). Two important attributes of these physiologic systems make them unlikely candidates for the “large effects—few genes” scenario. First, because these pathways are deeply embedded in normal physiology, they have been finely tuned by evolutionary selection and are controlled by many genes. Second, none of the specific phenotypes manifested as disease states would have been subject to positive selection because they occur with aging and have little impact on reproductive fitness. At the present time, strong evidence of selection in the human genome is restricted almost exclusively to loci that protect against epidemic infectious processes of relatively recent origin, with malaria still the best-

---

**Figure 2.** Age- and gender-adjusted hypertension prevalence by country and race in the 1990s. (From Cooper et al. 2005; reprinted, with permission, from the author.)
described example, or the ability to extract nutrient value from particular foods such as milk. Selection is almost always required to drive large interpopulation differences in functional genetic variants.

The search for racial differences in genetic predisposition must therefore be framed by this evolutionary context. If genetic susceptibility reflects multiple mutations in many genes that have no impact on fitness, it is highly improbable on statistical grounds that large numbers of these mutations would aggregate in one geographic population. This is even less probable for Africans, who share the vast majority of all variation in our species. Thus, within Africa, the entire spectrum of many phenotypic traits is observed (e.g., among the tallest populations and the shortest populations or groups with intestinal enzymes that allow metabolism of milk products and those without).

Although the genetic architecture underlying blood pressure regulation is still very poorly understood, the physiology makes it clear that many genes must be involved. The instantaneous response of blood pressure to the metabolic demands of every organ is an essential requirement for fitness among all animals with a cardiovascular system. In addition to being highly redundant and finely tuned, blood pressure regulation is a coordinated effort of multiple pathways—the nervous system, the heart, the kidneys, and the hypothalamic-pituitary-adrenal axis, among others. We now have empirical evidence that height is likely to be influenced by 4–500 genetic loci (Lanktree et al. 2011); it would be surprising if any fewer accounted for the distribution of blood pressure in the population. Large studies on blood pressure have failed to identify loci that have major effects, and although this may be in part owing to the difficulty in characterizing “average” blood pressure of an individual, it shows that genes with large effects are not common in human populations. For example, a study of 200,000 participants documented only a handful of genetic loci that were significantly associated with hypertension or blood pressure (The International Collaboration for Blood Pressure Genome-Wide Association Studies 2011). A combination of the 29 markers that had the strongest association with blood pressure raised the risk of hypertension by only 23%. Given that the lifetime risk of hypertension in the United States approaches 85%, the genetic information offers little predictive value so far. If, as expected, the distribution of genes with susceptibility alleles, and the alleles themselves, also vary to some degree among populations, and interactions with the environment are important, it will be extremely difficult, perhaps impossible, in the foreseeable future to calculate some net “population risk score” for multigenic traits such as hypertension.

Based on the evidence now available from studies of traits including height, we can model the underlying distribution of genetic variants that are common in populations (common meaning occurring in >5% of affected individuals). As seen in the accompanying figure, occasional genetic variants have been found that influence the complex traits by as much as 35%; however, these are very few in number (see smoothed estimate of distribution of regression coefficients associated with minor alleles for susceptibility loci, shown with red and without [blue] power adjustment in Park et al. [2011]). The underlying distribution is most likely a truncated normal distribution (i.e., a very large number of mutations with small effects, most of which have still escaped detection). Assuming that this pattern is typical for most diseases and that these mutations occur at random, this evidence provides the best empirical support yet available for the argument that large shifts in risk by continental or other population grouping are unlikely.

At the same time, of course, exceptions have been observed. As noted previously, large differences among populations are described for mutations that are protective against epidemic infections. Two additional examples have now been found for chronic diseases that are more common in U.S. blacks than in U.S. whites. Prostate cancer occurs about 2.5 times as often in black men, and mutations that are associated with a substantial increase in risk are now known to be very common in populations of African—at least West African—descent (Chang et al.
As this literature matures, however, the role of specific variants appears to be becoming more and more complex. In addition, a relatively uncommon form of chronic renal failure—focal segmental glomerulosclerosis—is strongly associated with a mutation in a gene related to lipid metabolism (*ApoL1*), and one copy of this version of the gene likewise occurs in almost 90% of West Africans (Rosset et al. 2011). Some evidence suggests that positive selection may be operating for the risk alleles for both prostate cancer and kidney disease—again, potentially through their role in immunity—although not all of the pieces of the puzzle fit together. As noted above, it remains exceedingly difficult to define the “causal” mutations that influence multigenic traits, and recent evidence suggests that *ApoL1* may not even be the target gene after all (S. Winkler, pers. comm.). However, if the original explanations are substantiated, these conditions would represent examples of race-specific susceptibility to common disease. Of course, a much more common condition based on race-specific genetic factors is present among Europeans who now live in sunny climates or at low latitudes, namely, their light skin. Needless to say, a host of other traits and diseases of smaller public health burden—from cystic fibrosis to sarcoidosis—also shows large variation among geographic populations. Sorting through the balance among geographic populations will become possible as the causal mutations are identified. A rough estimate of the contribution of known genetic variants to the overall health disparity between U.S. blacks and whites—summarized by life expectancy—yields a figure with an upper bound of 3% (Kaufman et al. 2013).

**EMERGING FROM THE BELIEF IN RACE TO THE OBSERVATION OF DIFFERENCES**

The contemporary experiment in large-scale genotyping has had contradictory effects on the meaning attributed to race in biomedicine. One major effect, certainly in the everyday perception of clinicians, has been a reinscribing of race as an important determinant of health and disease. Paradoxically, however, the massive outpouring of new knowledge about interindividual genetic variation has disposed of the proposition that race categories can serve as a broad proxy for genetic effects—both known and unknown. Furthermore, in many large cosmopolitan urban centers, the historical boundaries of race have blurred substantially. In New York City, for example, 36% of the residents were foreign born in 2005, speaking 170–200 languages. For the two largest official U.S. “minority” populations, blacks and Hispanics, a continuous spectrum of ancestral background is observed (Fig. 3) (Tayo et al. 2011); using ancestry information genetic markers, the figure depicts the proportion of ancestry from Europe and Africa.
The tension between the social prejudice regarding race and reasonable scientific evidence has thus been transformed. Given the opportunity to gradually sketch out the function of the genome and document its consequences for disease states, we can now begin to identify population-specific factors and assess their relative importance. Returning to the theme of historical context, it is worth noting that this tension has been a constant feature of modern biology since its origins. When he arrived in Paris in the 1860s to pursue the study of biology, Anténor Firmin, a Haitian intellectual and subsequent prime minister, was confronted with claims by the authorities in the field that “it is certain that the different races today are absolutely separate” (Firmin 1885). These claims were supported by evidence that he also had good reason to doubt, such as “the union of a Negro and a white woman are very often sterile.” After a thorough study of the current state of knowledge, he published a book entitled The Equality of the Human Races using existing evidence to advance essentially the same argument outlined in this essay (Firmin 1885). As Firmin wrote, “Yes, human beings can and do differ by their physical traits or the color of their skin. Yet, they are equal in intelligence and thought. Only a long process of perversion of the spirit and very powerful influences on the minds of White people could have made them overlook a truth that is so obvious and natural that it requires no scientific proof.”

His hopeful prediction for the utility of biological science to improve the human condition is equally modern: “We still have a chance of eradicating this prejudice. . . In the process, we will succeed in lowering the pretensions of an incomplete and ill-conceived science which continues unconsciousely to validate the most hurtful errors through assertions that are as suspect as they are perverse.” We can hardly lay claim to a better conceptual understanding than Firmin’s, but with the advantage of molecular technology, we can now make progress toward a science that has matured by being rooted in empirical evidence. Variation among geographic populations is real, and study of its origins can yield important biological insights. But there are no categories of race that segment human populations, and there are no mysterious qualities “in the blood” that justify the belief in racial superiority.

REFERENCES


Firmin A. 1885. The equality of the human races (reprinted by the University of Illinois Press, Chicago, 2002).

The International Collaboration for Blood Pressure Genome-Wide Association Studies. 2011. Common polymorphisms impacting blood pressure and cardiovascular...


Race in Biological and Biomedical Research

Richard S. Cooper

Cold Spring Harb Perspect Med 2013; doi: 10.1101/cshperspect.a008573

Subject Collection  Human Variation

Personalized Medicine and Human Genetic Diversity

Yi-Fan Lu, David B. Goldstein, Misha Angrist, et al.

Race in Biological and Biomedical Research

Richard S. Cooper

For additional articles in this collection, see http://perspectivesinmedicine.cshlp.org/cgi/collection/

Copyright © 2013 Cold Spring Harbor Laboratory Press; all rights reserved