

REVIEW

Race, genes, and health—new wine in old bottles?

Richard S Cooper

Accepted 17 October 2002

Technological advances that make it possible to identify nucleotide variants on a large scale are having an enormous impact on many fields in biology. This incursion of genetics is even being felt in the previously remote fields of epidemiology and public health. Yet while measuring the molecular basis of genetic variation will undoubtedly alter our understanding of the biological world, the implications for public health are much less clear. One of the most contentious areas involves the role of genetics in shaping health patterns among populations. In the 'pre-genomic' era race was the surrogate for genetic effects at the population level, and it must follow that molecular research will now largely transform that agenda. Although we are at the earliest stages several critical issues have already emerged, and a burgeoning literature exists in both epidemiology and genetics on the application of the new molecular technology to the study of race and health. While many new and unexpected findings have begun to emerge, there is also tension between constraints which tend to keep this advancing field within the existing assumptions and the impulse to re-examine prior assumptions as empirical data lead us in unexpected directions.

The intersection of race, genes, and health maps an enormous territory and the comments here will by necessity be very selective. To begin with, there is the perpetually troublesome concept of race itself. A clear example of the tension between reaffirmation of tradition and transformation of biological concepts can be seen in the discussion on the value of race as a tool to categorize humans.^{1–5} Relying on a variety of sources of molecular evidence, some investigators now argue that an old idea—humans can be grouped into continental races—has taken on new significance in public health and medicine.¹ The basis for this conclusion appears to rest on evidence that 'genetic differentiation is greatest when defined on a continental basis' and that multiple studies demonstrate that individuals from these groups can be clustered using molecular markers.

The central message put forward by those championing this updated version of race, wherein molecular data indeed show that genetic variation is discontinuous, represents a minority

position, however. Templeton, for example, uses molecular to arrive at a very different conclusion:

Human 'races' are not distinct lineages, and this is not due to recent admixture; human 'races' are not and never were 'pure'. Instead, human evolution has been and is characterized by many locally differentiated populations coexisting at any given time, but with sufficient genetic contact to make all of humanity a single lineage sharing a common evolutionary fate.²

Templeton also argues persuasively that some criteria (presumably a minimum value of F_{st} , or the proportion of genetic variance between as compared to within populations), is required to define a race or sub-species, and that all geographical arrangements are arbitrary. Although he did not reason from molecular evidence, Darwin came to the same conclusion on race:

Every naturalist who has had the misfortune to undertake the description of a group of highly varying organisms has encountered cases precisely like that of man; and if of a cautious disposition, he will end by uniting all the forms which graduate into each other ... ; for he will say to himself that he has no right to give names to objects which he cannot define.³

It is even far from clear that the important questions have been asked, or adequately tested thus far. First, one might wonder, why do we even want to know if races exist? Is this some sort of historical relic from the age when classification was the main object of nature science? But in reality we have passed those questions and must respond to the current agenda. To determine if broad racial groups exist, rather than asking, 'Can geographically distant groups be clustered?', it would seem more relevant to examine whether the global distribution of genetic variation can be divided into categories that correspond to continents. Addressing that question in 32 populations Romualdi *et al.* also come to the opposite conclusion: '... (T)here is little evidence, if any, of a clear subdivision of humans into biologically defined groups.'⁴ Of course, the individual data set as well as the analysis used by Romualdi *et al.* is subject to criticism, and they are likely to be quickly superseded by a much larger project. But in the end much of this debate seems to hang on

Correspondence: Dr RS Cooper, Department of Preventive Medicine and Epidemiology, Loyola University Stritch School of Medicine, 2160 S First Ave, Maywood, IL 60153, USA. E-mail: rcooper@lumc.edu

the interpretation of phrases like 'little evidence' and 'correctly classify', and even if we accept some group-level genetic difference and some degree of clustering that by itself provides little help in determining whether races exist. Darwin's caution still applies—what is missing is a coherent construct. However one chooses to interpret the data it is hard to see how a correct estimate for the number of human races, or the designation of what those races are, can be obtained without a definition. It is perhaps characteristic of the concept of race that it has retained its vitality in the scientific literature during the molecular age while being excluded from the requirement of a definition that meets the usual statistical requirements in epidemiology. From the point of view of the population genetics it is hard to see how progress will be made until this issue is resolved.

On the other hand, there is a contrary view that race has little or no biological meaning.^{5,6} Unless qualified, however, these views can be taken as inconsistent with the evidence, and have been an irritant to geneticists who see the importance of population variation in an array of conditions. One of the crucial distinctions in this debate relates to the size of the population being considered. Geographically localized populations clearly have different frequencies for many disease-related genetic variants. The primary point of contention, however, lies in whether it is useful to consider 'continental populations' as the most informative unit, or whether the descriptive value lies in groupings of the size of Corsicans, Scandinavians or Yorubans. Like most other measures of human variation, population size is obviously quantitative. Nonetheless, it has been well established that population of origin does provide useful information about the distribution of a number of genetic conditions. As suggested above by Romualdi *et al.*, however, the relationship between populations and prevalent genetic conditions correlates poorly with continental race. Fortunately this controversy at least is amenable to empirical study and a numerical solution should be forthcoming, while opinions will continue to diverge about the importance of the result.

In the absence of a statistical definition, however, one might ask, 'How does race matter in public health or medicine?' For continental race to have important biological meaning it must be shown to influence the broad public health patterns. This line of argument leads to a second crucial distinction in the race–health controversy—the nature of the conditions being examined. Most known genetic variants that are disease-associated are random mutations in sub-populations or result from specific regional selection; these are clearly not racial traits, in the continental sense. This result occurs because most monogenic disorders are caused by mutations of recent origin and they are therefore restricted to a smaller breeding population than represented by continents. Whether either regional selection or random mutations account for susceptibility to the traits that underlie common disease, to which the entire species is susceptible in varying degrees (e.g. regulation of body fat stores, maintenance of salt, water and blood pressure homeostasis, DNA repair, etc.), is unproven.

Acknowledging that one could define 'chronic disease' in a variety of ways, the racial patterns of the principal chronic conditions of public health interest are summarized in the Table. Despite widespread belief to the contrary, there is as yet no evidence that this pattern is shaped to any substantial degree by genetics. As acknowledged by even those in favour of the

Table 1 Health Status Measures in Racial/Ethnic Groups in the US, 1998

Cause of Death	Age-Adjusted Death Rates*			
	White	Black	Hispanic	Asian
All Causes	450.4	690.9	432.8	264.6
Heart disease	121.9	183.3	84.2	67.4
Coronary heart disease	79.2	92.5	54.7	42.9
Stroke	23.3	41.4	19.0	22.7
Cancer	121.0	161.2	76.1	74.8
COPD	21.9	17.7	8.5	7.4
Pneumonia/Influenza	12.7	17.4	9.8	10.3
Liver disease/Cirrhosis	7.1	8.0	11.7	2.4
Diabetes mellitus	12.0	28.8	18.4	8.7
HIV infection	2.6	20.6	6.2	0.8
External causes	46.7	68.8	44.7	24.4
Infant Mortality (/1000)	6.0	13.6	5.8	5.5
Life Expectance (years from birth)	77.3	71.3	(>80?)+	(>80)+

* Per 100 000

+ Estimate uncertain

Data sources; National Centre for Health Statistics

Health, United States, 2000 With Adolescent Health Chartbook, Hyattsville, Maryland: 2000.

biological importance of race, few genes underlying susceptibility to common diseases have been identified. One wonders, how, then, can one know that racial differences in common disease are genetic if one does not know what those susceptibility factors are? While there are many diseases with well-documented predisposing genes that vary among populations, they do not include complex chronic diseases.

In fact there are still enormous gaps in our understanding of the mechanism by which genetic variation influences physiological traits, like blood pressure, and associated common complex conditions, like hypertension. At the most basic level it should be pointed out that mutations that underlie essentially all monogenic disorders lead to altered (i.e. defective) proteins. These mutations are known as 'coding single nucleotide polymorphisms', or cSNP. Large scale genomic surveys have documented that cSNP are rare. Rare SNP are also of recent origin and, since they occurred after the original African diaspora, vary among population groups.⁷ Some observers have argued that a genetic basis for 'racial' differences is plausible since cSNP are rare, and therefore vary among groups.¹ However, while cSNP are the most likely candidates for variants affecting quantitative traits, that has not been widely demonstrated and the role of non-transcribed regulatory elements is still inadequately understood. It is well documented that quantitative traits in *Drosophila*, and less clearly in humans, are associated with common sequence variants in non-coding regions,^{8–11} although the meaning of this result is obviously still obscure. In fact, many observers think it likely that common variants underlie susceptibility to common disease, and since these diseases are observed worldwide, these alleles will be 'pan-ethnic'. Defining these mechanisms will require much deeper knowledge about genetic effects on common disease. Clearly an understanding of molecular mechanisms would also inform studies of racial/ethnic variation in disease. Persisting in efforts to fit our current notions of 'racial susceptibility' onto the limited evidence we have to date is unlikely to be very helpful, however.

Like all dramatic new technologies, genomics is promoted with substantial hype. In some quarters the molecular revolution even seems to have resulted in a hurriedly arranged truce between genetics and public health.^{12,13} Both sides have put down their weapons and run across no-man's land to fraternize and celebrate. But some of this jubilation may be premature; at the very least the declarations of a blissful future of mutual benefit are extreme. Up till now medical genetics has confined itself to rare diseases and diagnostic testing.^{12,13} Whether there is an important role for molecular genetics in diseases of public health significance is far from certain, and if race is the surrogate for genes we have even bigger problems. In contradistinction to ethnic or geographical sub-populations, continental race matters in public health precisely because it is associated with enormous variation in common diseases. But to conclude that we already know that there is a genetic dimension to race that matters because it confers susceptibility to the common chronic diseases is premature. The competing hypothesis, in support of which we have enormous evidence, is that this variation results from differences in environmental exposures. While geneticists rightly object to statements that no biological differences exist among populations, to rush to the opposite extreme and advance claims about the genetic underpinnings of chronic diseases having emerged under conditions of the modern industrial lifestyle is equally unhelpful.

The details of the technical arguments are ultimately not the greatest source of concern in this debate. As genomics enters the realm of public health, not only will changes be required in the study design and the inferences that follow, but the nature of the discourse surrounding those inferences must also change.^{12,13} To argue that the existence of the 'races of man' has been confirmed because of cluster analysis and description of superficial physical differences represents an intervention into public health that remains inadequately justified. Molecular data confers important analytical power, but it does not give population geneticists special authority to ignore historical context. Race is and always has been in essence a social construct. And I do not mean that some race effects are a result of social forces, but rather that the meaning of the word race is defined by its use. The construct of race simplifies and exaggerates the role of heritable factors.¹⁴ For geneticists to rush into this battleground, proclaiming to have verified by statistical analysis the popular wisdom that races exist is an inauspicious beginning to the molecular revolution in public health. This is not a question of hiding unhappy truths, but rather using scientific ideas for their social purpose.

Having argued that we should maintain an open mind in order to make use of what is new and unanticipated in the

emerging molecular data, and that strong conclusions are at best premature, let me offer my own (tentative) conclusions. The concept of race is unlikely to have value in public health—not only do the useful population units occur at a smaller size, race drags with it the dead weight of its noxious past. Populations will vary in the nature of their genetic risk profile, but this is unlikely to be true to any important degree for the common complex conditions. On that score we are much more alike rather than different. It will at any rate be a long time before we can test whether genes contribute much if anything to the observed public health disparities among demographic groups, however defined.

References

- ¹ Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. *Genome Biol* 2002;**3**:77 comment2007.1–2007.12.
- ² Templeton AR. Human races: A genetic and evolutionary perspective. *American Anthropologist* 1999;**100**:632–50.
- ³ Darwin C. *The Descent of Man*. Princeton, NY: Princeton University Press, 1981.
- ⁴ Romualdi C, Balding D, Nasidze IS *et al*. Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. *Genome Res* 2002;**12**:602–12.
- ⁵ Schwartz RS. Racial profiling in medical research. *N Engl J Med* 2001;**344**:1392–93.
- ⁶ Anon. Genes, drugs and race. *Nat Genet* 2001;**29**:239–40.
- ⁷ Tishkoff SA, Williams SM. Genetic analysis of African populations: human evolution and complex disease. *Nat Rev Genet* 2002;**3**:611–21.
- ⁸ Stam LF, Laurie CC. Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* 1996;**144**:1559–64.
- ⁹ Van Eerdewegh P, Little RD, Dupuis J *et al*. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002;**418**:426–30.
- ¹⁰ Mackay TFC. Quantitative trait loci in *Drosophila*. *Nature Rev Genet* 2001;**2**:11–20.
- ¹¹ Cox R, Bouzekri N, Martin S *et al*. Angiotensin-1 converting enzyme (ACE) plasma concentration in influenced by multiple ACE-linked quantitative trait nucleotides. *Hum Mol Genet*, in review.
- ¹² Anon. The HGP: end of phase 1. *Nat Genet* 2002;**30**:125.
- ¹³ Report of the Advisory Committee on Health Research. *Genomics and World Health*. Geneva: World Health Organization, 2002.
- ¹⁴ Marks J. Folk heredity. In: Fish JM (ed.). *Race and Intelligence. Separating Science from Myth*. New York: Lawrence Erlbaum Associates, 2002, pp. 95–112.

Commentary: Race, genetics, and disease— in search of a middle ground

Andrew J Karter

In the article, 'Race, genes, and health—new wine in old bottles?' in this issue of the *International Journal of Epidemiology*, Dr Cooper¹ presents an interesting perspective about the role of race in the study of genetics and disease. That paper is largely a critique of a recent paper in *Genome Biology* in which Risch *et al.*² reported evidence supporting categorization based on five major categories of self-identified race (Africans, Caucasians, Pacific Islanders, Asians, and Native Americans) and suggested that identifying genetic differences between these groups was 'scientifically appropriate'. Cooper counters that there is insufficient evidence that race or 'continental ancestry' has a biological (genetic) significance. He suggests that race is too crude a measure to have value in public health, and that finer-grained (sub-) categorizations (e.g. Scandinavians or Bantus) would be more informative than continental ancestry. He predicts that although populations show racial variation in genetic risk of rare, Mendelian diseases (e.g. Tay Sachs, sickle cell, thalassaemia), susceptibility allele distributions for the common, complex conditions that have world-wide distribution (e.g. diabetes or heart disease) are more likely 'pan-ethnic'. Therefore, Cooper believes that racial differences for common, complex diseases do not have a genetic origin. He argues that evaluating the role of race in genomic research is premature given the current lack of understanding of the genetic susceptibility for common complex disorders. Cooper concludes 'the concept of race is unlikely to have value in public health'.

In recent years, there has been a flurry of papers debating issues surrounding race, genetics, and disease, with Cooper being a major contributor. Unfortunately, much of the controversy raised in this paper stems from misunderstanding rather than true disagreement. Cooper cites the 'hype' surrounding the genomics era and the tendency of geneticists to view the world from a purely genetic dimension, ignoring the role of environmental factors. However, almost all genetic epidemiologists, Risch *et al.* included, would fully agree that environmental factors are important determinants of disease. While social epidemiologists have justifiably criticized molecular scientists espousing 'genetic-determinism', many social epidemiologists have promoted the equally unsubstantiated perspective that dismisses the influence of genetics on racial disparities in disease. Indeed, proposals from both extremes lack scientific merit. As the genomics era unfolds, the opportunity to test empirically whether genetics is a relevant factor in determining racial differences in complex diseases has arrived. The challenge will be to find the reasonable middle road that leads to appropriate research methodologies, a better understanding of disease aetiology, and practical public health interventions.³

Cooper's paper (as well as numerous previous papers)^{4–10} focuses on the processes of naming races and the appropriateness (scientific as well as political) of such naming (whether to 'lump or split') given the complex, multi-dimensional construct that is race. The variable 'race' has received special attention, largely because of society's deplorable history of racism and eugenics. However, from a purely scientific and methodological standpoint, this special attention is out of proportion. Risch *et al.* were simply posing a practical methodological solution to a vexing design problem in genetic studies of disease: appropriate stratification under the suspicion of racial differences in allelic distributions across populations. Note that I use the term 'stratification' in the classic epidemiological sense to mean analyses conducted separately in sub-groups (not to be confused with the term 'population stratification' used by population geneticists to mean statistical adjustment).

Stratified analyses are indicated when significant interactions ('effect modification', e.g. racial differences in the effect of alleles) are found in preliminary studies. While easily handled in the analysis, confounding (e.g. racial differences in the distribution of alleles, each of which have similar effect size across races) can also be minimized by studying restricted populations. It may be impossible to distinguish between effect modification and confounding by race with regards to the genetic effect before susceptibility alleles have been identified in non-stratified samples. However, since stratified analyses handle both situations, it seems the optimal solution. Additionally, since races are often not uniformly distributed in population-based samples, *post-hoc* strata-specific analyses may lack sufficient statistical power for the minority groups. Stratified sampling is therefore frequently designed into a study based on prior expectations. There exists adequate evidence of racial disparities to warrant stratification for genomic studies of many complex diseases. There are numerous examples of such racial disparities in disease (e.g. prostate cancer,¹¹ glaucoma,¹² chronic kidney disease,¹³ and diabetes-related lower extremity amputation)¹⁴ and response to therapy (e.g. high-dose interferon treatment for chronic hepatitis C)¹⁵ that persist after accounting for a wide range of potential confounders (e.g. access to care, socioeconomic status, health behaviours). Such residual race effects should at the very least raise some suspicion regarding the distribution and effect size of susceptibility alleles associated with these and other complex disorders. If we fail to detect genetic differences across races, then the stratified sampling will have no detrimental impact and pooled analyses will be fully appropriate.

A group of genetic scientists have suggested ignoring race in genetic studies.^{16,17} Justification for the 'race-neutral approach' is typically based on Lander's '99.9% identical rule'.¹⁸ However, these scientists often fail to mention that Lander also contends that although individuals are genotypically almost identical, the

tenth of a per cent of the genome that is different translates into roughly 3 million sequence differences, with some conferring dramatically differing risk of disease. For example, a single base pair change can cause haemochromatosis. By failing to design studies to accommodate the contingencies for interactions between race and genes, important racial differences in genetic susceptibility, if they exist, would likely remain undetected.

Cooper and others⁸ have justifiably questioned the analytical framework that proposes the utility of residual effects due to 'race' as a proxy for genetics (after adjusting for expected alternative explanatory variables, e.g. socioeconomic status). Fixed attributes such as 'race' fail in the context of causal modelling because they are not substitutable (see discussions of the counterfactual model of causality)¹⁹ and do not yield the answers that we actually seek (e.g. what would the health of individuals of race X be given they experienced life as race Y?). Because of these caveats, testing population-specific hypotheses (e.g. race specific allele–disease associations) makes more sense than making inferences based on residual effects. Clearly, for scientists who have evidence suggesting potential racial differences in disease susceptibility, the only way to accomplish this is to use stratified (by race) designs.

The looming question is how to operationalize population-stratification. Cooper objects to the use of self-identified race because of its lack of precise definition. Risch *et al.* suggest that self-identified race is an adequate approach to human categorization in biomedical and genetic research, but do not deny that it is a very crude categorization. These authors simply suggest that genetic differentiation, including disease susceptibility alleles, although certainly also occurring on a finer scale, is greatest when defined by these five major racial categories. Moreover, finer-scaled categorization is quite impractical from a research operations standpoint.

Instead of relying on self-reported race, many scientists, including the leadership of the Human Genome Project, support using race-specific markers (e.g. microsatellite markers) to facilitate population stratification. However, marker-based studies would require genotyping *before* population stratification could take place, demanding expensive data collection and laboratory assays on inflated sample frames to accommodate the identification of sufficient individuals in minority populations. Moreover, Risch *et al.* suggest that the confluence of social, cultural, behavioural, and environmental variables that are associated with self-identified race introduce confounding between genetic and environmental risk in an ethnically heterogeneous, race-neutral study. By ignoring self-identified race, or even when stratifying on race specific genetic markers, environmental culprits may be missed due to our inability to disentangle the residual effects of confounding. If disease variation was due to racially varying cultural practices, then self-identified race would be a better adjuster than genetic markers given that cultural practices would not be maintained over time if its members could not identify each other.^{20,21} The use of self-identified race provides the most practical resolutions for problems of confounding (e.g. matching, adjusting, or restriction) and effect modification (stratification).

Epidemiologists have a long history of benefiting from designs that stratify samples in rather crude categories that are surrogates for ill-defined constructs (e.g. social position indicators such as income, education, and occupation). Despite the

imperfection of such measures, they have frequently moved us closer to understanding the determinants of disease. Similarly, self-identified 'race', albeit a crude proxy, may accomplish the same in the context of genomic research. Many, if not most, scientific hypotheses that are tested and later shown to be important findings were stimulated not by *a priori* hypotheses and well-formed theory, but rather by exploration of data. I suggest that our abilities to move epidemiological and genetic research forward would be greatly impeded if we were to avoid the use of surrogates such as self-identified race.

Cooper correctly points out the 'enormous gaps in our understanding of the mechanism by which genetic variation influences common complex conditions'. However, the fact that genomic research has made relatively little progress in untangling complex genetic disorders is not grounds for ignoring race. In fact, our progress may have been hampered due to the lack of adequate stratification to date in the face of differing effect size or distribution of susceptibility alleles across populations. Cooper should embrace the use of self-identified race as a design variable, as it may be the only practical way that we may fill those large gaps in our understanding. There is no denying that epidemiological studies have demonstrated relevant racial variation in the incidence, prevalence, and natural history of many diseases. I concur with Cooper's suggestion that the dominant cause is likely environment (social, cultural, economic, political) in the US where, unlike countries with socialized medicine, access and quality of health care varies greatly by race^{22,23} and socioeconomic position. In general, the poorer US minority populations experience reduced access to and quality of health care. Moreover, there is evidence that internalized racism impacts health,^{24,25} and physicians' perceptions of and interactions with patients may differ by race.²⁶

However, these assertions in no way preclude the possibility that complex diseases may also differ by race due to genetics. I agree with Cooper's assertion that these effects may be attributable to residual confounding by unspecified environmental factors. However, if Cooper's allegation is true that geneticists are 'using scientific ideas for their social purpose', it is certainly more the exception than the rule. In fact, by avoiding a full exploration of the potential genetic contribution because of the social baggage that accompanies the concept of 'race' (however it may be measured), we may do an even greater disservice to those at-risk populations we are trying to protect. Indeed Cooper himself seems also to support a genetic approach to racial differences in a recent paper²⁷ on admixture linkage disequilibrium mapping, the primary assumption of which is that racial differences in prevalence are due to allele frequency differences at genetic loci.

Cooper must be congratulated for his meaningful contributions, present and past, to this important public health debate. It is clear that both geneticists and epidemiologists have a similar goal in mind, which is to understand causes of disease in general, and in so doing, reduce racial disparities. He predicts that identifying race-specific genetic aetiologies will not have a significant impact on public health. Although environmental risk factors (e.g. behavioural, social, cultural, political, economic) may be a bigger culprit than genetics with regard to racial disparities in disease, identifying genetic aetiologies with even small attributable risks will further our understanding of the tangled web of aetiological pathways involved in complex

disease. If this understanding leads to identification of novel pharmacological interventions, then the impact of our genetics (i.e. gene products) may be in a sense 'modifiable'. Paradoxically, while environmental factors are typically thought of as modifiable, they may ultimately prove less yielding to change than our biology. Regardless, epidemiologists cannot afford to ignore either genes or environment.

References

- ¹ Cooper RS. Race, genes, and health—new wine in old bottles? *Int J Epidemiol* 2003;**32**:23–25.
- ² Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. *Genome Biol* 2002;**3**:COMMENT2007.
- ³ Lin SS, Kelsey JL. Use of race and ethnicity in epidemiologic research: concepts, methodological issues, and suggestions for research. *Epidemiol Rev* 2000;**22**:187–202.
- ⁴ Cooper R. A note on the biologic concept of race and its application in epidemiologic research. *Am Heart J* 1984;**108**:715–22.
- ⁵ Cooper R, David R. The biological concept of race and its application to public health and epidemiology. *J Health Polit Policy Law* 1986;**11**:97–116.
- ⁶ Cooper R, Steinhauer M, Miller W, David R, Schatzkin A. Racism, society, and disease: an exploration of the social and biological mechanisms of differential mortality. *Int J Health Serv* 1981;**11**:389–414.
- ⁷ Cooper RS. Health and the social status of blacks in the United States. *Ann Epidemiol* 1993;**3**:137–44.
- ⁸ Kaufman JS, Cooper RS. Commentary: considerations for use of racial/ethnic classification in etiologic research. *Am J Epidemiol* 2001;**154**:291–98.
- ⁹ Kaufman JS, Cooper RS. Seeking causal explanations in social epidemiology. *Am J Epidemiol* 1999;**150**:113–20.
- ¹⁰ Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology* 1997;**8**:621–28.
- ¹¹ Shibata A, Whittemore AS. Genetic predisposition to prostate cancer: possible explanations for ethnic differences in risk. *Prostate* 1997;**32**:65–72.
- ¹² Tielsch JM, Sommer A, Katz J, Royall RM, Quigley HA, Javitt J. Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. *JAMA* 1991;**266**:369–74.
- ¹³ Tarver-Carr ME, Powe NR, Eberhardt MS *et al.* Excess risk of chronic kidney disease among African-American versus white subjects in the United States: a population-based study of potential explanatory factors. *J Am Soc Nephrol* 2002;**13**:2363–70.
- ¹⁴ Karter AJ, Ferrara A, Liu JY, Moffet HH, Ackerson LM, Selby JV. Ethnic disparities in diabetic complications in an insured population. *JAMA* 2002;**287**:2519–27.
- ¹⁵ De Maria N, Colantoni A, Idilman R, Friedlander L, Harig J, Van Thiel DH. Impaired response to high-dose interferon treatment in African-Americans with chronic hepatitis C. *Hepatogastroenterology* 2002;**49**:788–92.
- ¹⁶ Schwartz RS. Racial profiling in medical research. *N Engl J Med* 2001;**344**:1392–93.
- ¹⁷ Wilson JF, Weale ME, Smith AC *et al.* Population genetic structure of variable drug response. *Nat Genet* 2001;**29**:265–69.
- ¹⁸ Lander ES. Mapping heredity: using probabilistic models and algorithms to map genes and genomes. In: Lander ES, Waterman MS (eds). *Calculating the Secrets of Life, Applications of the Mathematical Sciences in Molecular Biology*. Washington, DC: National Academy Press, 1995, pp. 25–55.
- ¹⁹ Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol* 2002;**31**:422–29.
- ²⁰ Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev* 2002;**11**:513–20.
- ²¹ Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev* 2002;**11**:505–12.
- ²² Schneider EC, Zaslavsky AM, Epstein AM. Racial disparities in the quality of care for enrollees in medicare managed care. *JAMA* 2002;**287**:1288–94.
- ²³ Smedley BD, Stith AY, Nelson AR. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Institute of Medicine, 2002. Washington, DC: National Academy Press.
- ²⁴ Krieger N, Sidney S, Coakley E. Racial discrimination and skin color in the CARDIA study: implications for public health research. Coronary Artery Risk Development in Young Adults. *Am J Public Health* 1998;**88**:1308–13.
- ²⁵ Tull ES, Chambers EC. Internalized racism is associated with glucose intolerance among Black Americans in the U.S. Virgin Islands. *Diabetes Care* 2001;**24**:1498.
- ²⁶ van Ryn M, Burke J. The effect of patient race and socio-economic status on physicians' perceptions of patients. *Soc Sci Med* 2000;**50**:813–28.
- ²⁷ Collins-Schramm HE, Phillips CM, Operario DJ *et al.* Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *Am J Hum Genet* 2002;**70**:737–50.