STATE OF THE ART AND SCIENCE
Complex Systems for a Complex Issue: Race in Health Research
Abdulrahman M. El-Sayed, MD, DPhil

Population health research has, as its main aim, characterizing the causes and consequences of health and disease. The process of rigorously testing hypotheses about what influences population health is challenging, particularly when considering the health effects of inalterable or potentially deleterious exposures. While the design, execution, and interpretation of population health experiments is itself arduous, trials are only even possible when the health effects of alterable, potentially beneficial treatments are being interrogated. By contrast, analyzing the effects of inalterable, potentially deleterious exposures is substantially more problematic. Absent the luxury of clinical trials, and given obvious ethical or logistical limitations, the enterprise has largely been limited to the analysis of large datasets derived by following individuals over time to attempt to identify the effects of these exposures.

Race is one such "inalterable exposure," the rigorous study of which is methodologically complicated. Race is doubtless among the most important predictors of health in heterogeneous societies such as the United States, associated with myriad health metrics. Race is an important predictor of birth weight, gestational age, and risk of infant mortality [1]; it is associated with health behaviors that may influence health outcomes throughout life, such as physical activity [2] and health services access and use [3]; finally, race is a robust predictor of the timing and nature of the end of life, associated with age-specific, cause-specific, and all-cause mortality [4].

It is similarly clear, however, that race is a poorly defined and nonspecific term [5]. It is but a proxy for a number of other (possibly causal) factors that may relate to health. The growing literature that has identified consistent associations between race and health outcomes sheds little light on the mechanisms that produce those differences [1-5]. There are a number of plausible hypotheses. Biological scientists point to concrete differences in the observable phenotypes that differentiate racial groups and argue that there must be genotypic differences underlying these observed differences and, therefore, that there may also be genotypic differences underlying differences in health outcomes [6]. However, genetic variation within racial groups is substantially greater than it is between such groups [7], and the minimal variation between groups has not been shown to have substantial relation with health outcomes [8]. Nevertheless, potential differences in the distribution of risk alleles or the epigenetic modification of alleles by race remains a viable hypothesis for racial differences in certain health outcomes [6].

By contrast, social epidemiologists contend that race shapes health by way of a number of social mechanisms [9-11]. First, variation in the experiences of discrimination and xenophobia both between and among historically marginalized groups is predictive of differences in health outcomes among them [9]. For example, Geronimus observed that black-white disparities in birth weight increased with increasing maternal age at parturition [12, 13]. She reasoned this "weathering" of black relative to non-Hispanic white mothers represented the cumulative effect of lifelong exposure to a caustic social environment experienced by black, but not non-Hispanic white, mothers as they aged [12, 13]. Race also has implications for the social environments to which people are exposed, the ways resources are allocated, and ultimately, the health experiences of members of each group. Supporting these arguments in the health literature are consistent correlations between race, socioeconomic position, and health outcomes, such that marginalized groups are persistently disadvantaged and suffer the health consequences thereof [14-16].

Understanding the mechanisms that underlie the role of race in shaping population health is particularly challenging. How do we, for example, disaggregate the effects of discrimination from those of individual and contextual socioeconomic disadvantage when considering the etiology of racial differences in health? This is especially challenging because these factors are mutually reinforcing: discrimination has implications for educational attainment and income prospects throughout the life course, which shape the neighborhoods and locales—the environments—into which individuals of different racial groups select, which ultimately feed back into their social experiences [9, 17].

As inherited traits are fundamentally inalterable, we are left with observational epidemiologic studies to attempt to isolate the effects of these influences. However, our limited methodology, founded upon regression modeling, forces us to make increasingly less tenable assumptions to that end [18, 19]. For example, traditional health research approaches are beholden to the assumption that individuals are autonomous agents whose interactions with others do not influence the individual agent's outcomes. But what is discrimination if not a process resulting from human interaction—the influence of one agent's actions on another? Similarly limiting is the incapacity of observational approaches to efficiently represent context [20], conflating aggregate characteristics of the individuals residing in particular spaces with the environmental influences of those spaces themselves [21]. For example, studies demonstrating associations between neighborhood poverty and poor health cannot accurately differentiate between the influences of these neighborhoods and the influences of the characteristics of the individuals living therein [21]. This limitation prevents us from teasing apart the health implications of individual and context-level socioeconomic deprivation on the health of people in marginalized groups. Lastly, because observational studies only represent levels of exposure and outcomes at specified points in time, they cannot deal effectively with the reciprocity or feedback between exposures or between exposures and outcomes that we understand to be important in shaping racial disparities in health [18, 19].

Complex systems approaches may be particularly important for understanding the mechanisms underlying the influence of race on population health [18, 19]. Whereas traditional epidemiologic approaches are designed to test hypotheses about population-level associations between exposures and outcomes of interest, systems approaches are designed to test hypotheses about the dynamics that underlie the data. They force investigators to represent and test hypotheses about the mechanisms that produce population health and disease while freeing them of some of the problematic assumptions imbedded in regression models. By embracing, rather than ignoring, important factors such as inter-individual interactions, environmental influences, and reciprocity and feedback, these tools enable us to understand why race is so strong a predictor of differences in health.

One particularly promising systems method for health research about race is agent-based modeling [19]. Agent-based models simulate individuals with various attributes and preferences who interact with one another via dynamic social networks and who are situated in particular contexts to which they contribute and that influence their behaviors in turn. Agent-based models can account for an individual with her particular characteristics, the neighborhood in which she lives, as well as the network of friends and family she has—all changing dynamically with time, according to the joint behaviors of the agents in the model. Hence, these models allow investigators to consider the population-level consequences of individual-level behaviors, and the population-level dynamics they model emerge from the behaviors of the individuals the population comprises. Agent-based models have already demonstrated their utility in health research. Investigators have employed them to demonstrate the social mechanisms that underlie differences in such activities as walking behavior in an urban environment and to analyze the value of using social networks to mitigate obesity [22, 23].

Agent-based models hold particular promise for the study of race and health because they allow investigators to explicitly articulate, represent, and, therefore, test the social and biological mechanisms hypothesized to underlie influences of race on health. Rather than continuing to rely on race as a nonspecific proxy for a number of underlying exposures of interest, agent-based models allow investigators to represent each of these underlying exposures explicitly as aspects of agent attributes or behaviors. From socioeconomic characteristics and experiences of personalized or structural discrimination to the locus-specific alleles that shape genotypes, agent-based models allow investigators to deconstruct race to its core components to better understand each mechanism contributing to the systematic racial differences in health that we so routinely observe.

While agent-based modeling and complex systems approaches are fledgling methodologies in health research, they are well suited to characterizing the myriad influences of race on health. The benefits of these approaches are clear for investigators and policymakers alike. They allow investigators a more granular, hypothesis-based way to understand and disentangle the mechanisms that shape the racial differences we observe by enabling us to bypass several obfuscating

assumptions that limit our inference regarding the role of race in health. By deconstructing race in health research, we may better understand the dynamics that operate to produce race-related disparities in health and how mechanisms such as interactions, context, and feedback may be shaping them. For these reasons, complex systems approaches have important implications for policy as well. Equipped with a more comprehensive yet focused understanding of the dynamics underlying racial disparities, these approaches may provide policymakers interested in tackling these inequalities with better defined, more relevant levers on which to focus their efforts.

References

1. Spong CY, Iams J, Goldenberg R, Hauck FR, Willinger M. Disparities in perinatal medicine: preterm birth, stillbirth, and infant mortality. *Obstet Gynecol.* 2011;117(4):948-955.
2. Crespo CJ, Smit E, Andersen RE, Carter-Pokras O, Ainsworth BE. Race/ethnicity, social class and their relation to physical inactivity during leisure time: results from the Third National Health and Nutrition Examination Survey, 1988–1994. *Am J Prev Med.* 2000;18(1):46-53.
3. Fiscella K, Franks P, Doescher MP, Saver BG. Disparities in health care by race, ethnicity, and language among the insured: findings from a national sample. *Med Care.* 2002;40(1):52-59.
4. Murray CJ, Kulkarni SC, Michaud C, et al. Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Med.* 2006;3(9):e260. Erratum in: *PLoS Med.* 2006;3(12):e545.
5. Manly JJ. Deconstructing race and ethnicity: implications for measurement of health outcomes. *Med Care.* 2006;44(11 Suppl 3):S10-S16.
6. Collins FS. What we do and don't know about 'race','ethnicity', genetics and health at the dawn of the genome era. *Nat Genet.* 2004;36(11 Suppl):S13-S15.
7. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science.* 2008;319(5866):1100-1104.
8. Lohmueller KE, Mauney MM, Reich D, Braverman JM. Variants associated with common disease are not unusually differentiated in frequency across populations. *Am J Human Genetics.* 2006;78(1):130-136.
9. Williams DR, Neighbors HW, Jackson JS. Racial/ethnic discrimination and health: findings from community studies. *Am J Pub Health.* 2003;93(2): 200-208.
10. Williams DR. Racial/ethnic variations in women's health: the social embeddedness of health. *Am J Public Health.* 2002;92(4):588-597.
11. Williams DR. Race, socioeconomic status, and health. The added effects of racism and discrimination. *Ann NY Acad Sci.* 1999;896(12):173.
12. Geronimus AT. The weathering hypothesis and the health of African-American women and infants: evidence and speculations. *Ethn Dis.* 1992;2(3):207-221.
13. Geronimus A. Black/white differences in the relationship of maternal age to birthweight: a population-based test of the weathering hypothesis. *Soc Sci Med.* 1996;42(4):589.
14. Wilson WJ. When work disappears: New implications for race and urban poverty in the global economy. Ethnic and racial studies 1999;22(3):479.

15. Ward TC, Mori N, Patrick TB, Madsen MK, Cisler RA. Influence of socioeconomic factors and race on birth outcomes in urban Milwaukee. *WMJ.* 2010;109(5):254-260.
16. Pickett K, Ahern J, Selvin S, Abrams B. Neighborhood socioeconomic status, maternal race and preterm delivery: a case-control study. *Ann Epidemiol.* 2002;12(6):410.
17. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. Public health reports 2001;116(5):404.
18. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. *Int J Epidemiol,* 2010;39(1):97.
19. El-Sayed AM, Scarborough P, Seemann L, Galea S. Social network analysis and agent-based modeling in social epidemiology. *Epidemiol Perspect Innov.* 2012;9(1):1.
20. Auchincloss AH, Diez Roux AV. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *Am J Epidem.* 2008;168(1):1-8.
21. Oakes JM. The (mis) estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med.* 2004;58(10):1929-1952.
22. Yang Y, Diez Roux AV, Auchincloss AH, Rodriguez DA, Brown DG. A spatial agent-based model for the simulation of adults' daily walking within a city. *Am J Prev Med.* 2011;40(3):353-361.
23. El-Sayed AM, Seemann L, Scarborough P, Galea S. Are network-based interventions a useful antiobesity strategy? an application of simulation models for causal inference in epidemiology. *Am J Epidemiol.* 2013;178(2):287-295.

Abdulrahman M. El-Sayed, MD, DPhil, is an assistant professor in the Department of Epidemiology at Columbia University in New York City. His research considers the social production of population health, health disparities, and systems methods for epidemiologic research.

Related in VM
Race, Discrimination, and Cardiovascular Disease, June 2014

Minority-Group Recruitment Goals in Federally Funded Clinical Research: What's in a Number? June 2014

Race: A Starting Place, June 2014

Pain and Ethnicity, May 2013

The viewpoints expressed on this site are those of the authors and do not necessarily reflect the views and policies of the AMA.