American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2012.

Vol. 175, No. 7
DOI: 10.1093/aje/kws138
Advance Access publication:
March 12, 2012

Commentary

Cardiovascular Epidemiology in a Changing World—Challenges to Investigators and the National Heart, Lung, and Blood Institute

Paul D. Sorlie, Diane E. Bild, and Michael S. Lauer*

* Correspondence to Dr. Michael S. Lauer, Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, 6701 Rockledge Drive, Room 8128, Bethesda, MD 20892 (e-mail: lauerm@nhlbi.nih.gov).

Initially submitted December 28, 2011; accepted for publication February 14, 2012.

Over the past 60 years, revolutionary discoveries made by epidemiologists have contributed to marked declines in cardiovascular disease morbidity and mortality. Now, in an era of increasingly constrained resources, researchers in cardiovascular epidemiology face a number of challenges that call for novel, paradigm-shifting approaches. In this paper, the authors pose to the community 4 critical questions: 1) How can we avoid wasting resources on studies that provide little incremental knowledge? 2) How can we assure that we direct our resources as economically as possible towards innovative science? 3) How can we be nimble, responding quickly to new opportunities? 4) How can we identify prospectively the most meritorious research questions? Senior program staff at the National Heart, Lung, and Blood Institute invite the epidemiology community to join them in an ongoing Web-based blog conversation so that together we might develop novel approaches that will facilitate the next generation of high-impact discoveries.

cardiovascular diseases; epidemiology; National Institutes of Health (U.S.); research

Abbreviations: AIDS, acquired immunodeficiency syndrome; NHLBI, National Heart, Lung, and Blood Institute; NIH, National Institutes of Health.

For more than 10 years, scientists were unable to describe the molecular structure of a protein-cutting enzyme from an AIDS-like virus found in rhesus monkeys (1, 2). Knowing the structure could help the researchers develop novel antiviral drugs. Stumped, the scientists presented their problem to players of ''Foldit,'' an online game that involves many players who team up and use 3-dimensional problem-solving skills. The gamers solved the problem in 10 days! The game designer stated, ''People have spatial reasoning skills, something computers are not good at. Games provide a framework for bringing together the strengths of computers and humans'' (2). Sometimes we need wholly new playbooks to solve complex problems.

As we progress through a new decade, one in which resources are likely to be increasingly constrained, we in cardiovascular epidemiology face a number of critical challenges that call for novel solutions. How can we avoid wasting resources on studies that provide little incremental knowledge? How can we assure that we direct our resources towards innovative science as economically as possible? How can we be nimble, responding quickly to new opportunities? And how can we identify prospectively the most meritorious research questions? Our crucial test will be how well we adopt new mindsets and move beyond conventional thinking, in order to formulate creative, viable, and valid solutions. In this commentary, we at the National Heart, Lung, and Blood Institute (NHLBI) begin to pose these challenges and enlist cardiovascular epidemiologists to join us in seeking solutions.

CHALLENGES

Challenge 1: avoiding studies that add little incremental knowledge or uncertain validity

Ten years ago in these pages, Kuller (3) admonished investigators to avoid redundant epidemiology. He criticized researchers who demonstrated (again) that weight gain is a risk factor for diabetes or who described associations between use of vitamins and clinical outcomes when definitive clinical trials were ongoing. Not only do redundant studies waste research

dollars and investigators' time, they exacerbate perceptions in the lay and scientific communities that epidemiology has little value. Kuller faulted peer review and the research funding process for being overly conservative, shying away from high-risk science. More recently, Ioannidis et al. (4) described the opposite problem—the abundance of new but likely false epidemiologic findings that far outweigh true findings. Some novel, first-report studies have been published without due diligence to multiple comparisons, use of cautious *P* values, or consideration of biologic plausibility. Later, when the findings are contradicted, many observers lose all confidence in epidemiology. Kuller and Ioannidis et al. highlight the consistently expressed perspectives of our harshest critics: The results are already known (in other words, "so what?") or they get refuted (in other words, even if there is a "so what," can we believe them?).

Challenge 2: assuring innovative science at a reasonable cost

How do we create and find innovative strategies that combine rigorous scientific methods with a meaningful impact on public health and scientific thinking? While we seek "the best science for an affordable cost" as an implicit driving paradigm, we are confronted with the problem that innovation often means higher costs for projects that carry higher risk of failure. We face a number of daunting cost drivers, including the high labor costs intrinsic to large, scientifically diverse research teams; the requirement of enormous sample sizes for some research questions; and increasingly expensive technology. Our dilemma is that we believe we need these resources to advance the science.

We in epidemiology have long appreciated the meaning and importance of "team science." For us, this is not a novel concept. Successful epidemiology teams include people with expertise in basic, behavioral, and social sciences; informaticists, methodologists, and statisticians; and clinicians who can understand where gaps in knowledge impede potential advances in health care. Modern-day epidemiologic teams' approaches also need experts in genetics, proteomics, metabolomics, noninvasive imaging technologies, and systems scientists. All-star teams come at a cost. Identifying and recruiting participants to assemble a representative sample, examining them for hours in a clinic, and performing sophisticated tests, often including imaging, are labor-intensive activities. Costs are compounded according to the number of participants included—typically in the thousands.

Challenge 3: responding rapidly to new opportunities

We need to adapt rapidly to new research avenues. We have a track record in epidemiology, having demonstrated the ability to use stored samples to efficiently evaluate new biomarkers. We created collaborative, highly productive, international research consortia to maximize sample size. However, scientists often find it difficult to incorporate new imaging and advanced data collection technologies into ongoing studies. Study planning and implementation takes many months to years, and the National Institutes of Health (NIH) grant application and peer-review system is not rapid. While the NIH has changed its peer-review system in recent years to improve efficiency (5, 6), the time from application submission to funding is generally never less than 9 months, and often is almost twice that long.

Epidemiology researchers also have difficulty enabling real-time responses to societal and policy changes. For example, immigration has substantially altered the composition of the American population, bringing new health problems related to adaptation to a new culture and differentials in health-care access and behaviors. Recent economic problems have led to greater unemployment, less health insurance, and widening health and economic disparities. Health-care reform brings a whole new set of policy uncertainties. Epidemiologists need strategies to evaluate these changes in a timely manner in order to characterize their impact on health and to identify possible solutions.

Challenge 4: identification of research questions with the greatest merit

Cardiovascular epidemiologists should focus on what is ultimately important for the health of the nation and world. So how do we identify the most important research questions? The usual short definition of epidemiology is "the study of the distribution and determinants of disease." Recognizing that we study health as well as disease, the key words for this commentary are "distribution" and "determinants." Studies of distribution must continue but will require innovative efforts to do much more with less. To study determinants, should epidemiology function like basic science, attempting to discover associations that lend insight into the mechanisms and natural history of disease, without worrying about whether there are practical implications for public health? While the results may enhance understanding of human biology, we don't know whether they will create hypotheses for paradigm shifts in preventive medicine and public health. Should we do discovery epidemiology and data-mining without defined hypotheses, or should we put more emphasis on epidemiologic studies that identify and describe targets for disease prevention? Is it truly worthwhile to continue trying to account for unexplained prediction of known risk factors when analyses generally show minimal added prediction (7, 8)?

We certainly want the best possible studies to establish the value of putative prevention strategies. The conflict, though, is that we might find certain genotypes, biomarkers, or imaging findings predictive of disease and consequently find ourselves incorrectly tempted to conclude that using these new indicators clinically will necessarily prevent disease. We might argue that we epidemiologists do ourselves no favor when we try to "have it both ways," claiming that the epidemiologic method is invaluable for learning about disease and generating hypotheses while suggesting that our results are directly actionable. While there should be an ultimate promise of health benefit from epidemiologic studies, the epidemiology community is challenged to determine what highest-priority research should be done and to define the potential in a meaningful way.

NHLBI STRATEGIES FOR ADDRESSING THESE CHALLENGES

To address these challenges in broad strokes, the NHLBI queried more than 600 scientists and synthesized their responses

in 2007 in *Shaping the Future of Research: A Strategic Plan for the National Heart, Lung, and Blood Institute* (9), a plan that still guides its scientific priorities. The NHLBI research spectrum spans a range from basic laboratory science to broad clinical effectiveness trials. The Strategic Plan envisions research as focusing on the transitions from form to function, function to causes, and causes to cures. Traditionally, epidemiology was not seen as contributing to the goals that are included in "form to function." In current practice, however, and as emphasized in the Strategic Plan, epidemiology plays a major role in all of the major goals. We now appreciate that population scientists greatly facilitated the now widely acclaimed discoveries of genetic variation. Epidemiologists have delineated the role of inflammation in disease. Epidemiologists generate hypotheses for clinical trials, establish the bases for personalized medicine, and evaluate health care and public health effectiveness.

In order to maintain the dynamism of our strategic planning, the NHLBI continually seeks informal and formal input from the research community. The Institute organizes dozens of scientific workshops annually to engage extramural scientists, who provide ideas, perspectives, and advice about scientific direction and possible initiatives. We seek to use workshops and working groups as a forum for explicit thought about how best to avoid low-value studies, to develop innovative methods, to respond rapidly to new opportunities, and to identify the most important research questions. For example, several recent workshops identified national needs relating to epidemiologic research, including recommendations for large-scale studies for genomics and proteomics with long follow-up to capture progression of disease and/or other health-state transitions (10); enhanced data collection on the natural history of atrial fibrillation through intensive monitoring of high-risk subsets of the population (11); building the scientific foundation for personal genomics (12); and developing a registry to enable study of sudden cardiac death in the young (13).

As a research community, we must develop strategies that will build on our current strengths, address challenges, and at the same time cost less. Like other sectors of the economy, whether in research or elsewhere, we should look upon current fiscal challenges as an opportunity to figure out how to do more (indeed much more) with less. Several strategies are already apparent and are worthy of serious consideration:

- Leverage and even consolidate existing infrastructures (14), including existing cohort studies within the NHLBI and other NIH institutes; large health-care databases from health maintenance organizations (15); the Veterans Administration or military; disease registries assembled by professional societies; and the Clinical and Translational Science Award programs.
- Determine how best to exploit electronic medical records with consideration of incorporating natural language processing while also recognizing their inherent limitations (16, 17).
- Decentralize examinations by seeing participants in their homes using standardized protocols (REGARDS (18)) and/or by employing efficient examination operations, as is done with the UK Biobank (http://www.ukbiobank.ac.uk), which successfully enrolled over 500,000 subjects at a total cost below $100 million.

NEW RESEARCH THEMES

Two new research themes in epidemiology that have generated much discussion are expansion of sample sizes from the thousands to the hundreds of thousands and incorporation of comparative effectiveness studies. Both of these themes have their share of design, methodological, and cost problems, but because of their potential to advance science and inform health care, they deserve careful discussion.

Perhaps spurred by the emergence of population genomics, epidemiologic studies often include enormous numbers of participants and patients. As researchers seek to learn more about the associations of genetic, epigenetic, proteomic, and environmental variations with disease, we are witnessing a new era of "mega-epidemiology" studies—based on consortia of existing cohort studies or new studies—with sample sizes in the hundreds of thousands to millions (19–21). These types of studies tend to fall into 1) analyses of clinical or administrative databases or 2) mergers or creation of huge cohorts with direct measures of clinical, biologic, and sociodemographic phenotypes—with some calls for combination of the 2 models. These approaches are not original. For example, in the United Kingdom "Million Women Study," investigators combined survey data with electronic health records to investigate the association of alcohol intake with cancer risk in nearly 1 million women (22).

Some researchers have successfully used huge administrative databases, such as those maintained by the United States Renal Data System (http://www.usrds.org) and the Kaiser Permanente health plan, to study population trends in various disease conditions, including end-stage renal disease (23), myocardial infarction (15), and outcomes research (24). Others promote electronic health records and registries as a platform for clinical epidemiology (14), for genomic analysis (25, 26), and for clinical registry trials (27). Some hospitals and health plans are collecting biospecimens from huge numbers of patients, with the idea that local but large biorepositories can be linked to electronic health data and to administrative claims data. While these approaches have been successfully used to identify large numbers of disease cases, accuracy for some diagnoses and limitations in the data available for analysis remain a concern (16, 17).

Another recent theme has been inclusion of observational studies in the spectrum of activities in the burgeoning field of comparative effectiveness research, which has been defined as "a rigorous evaluation of the impact of different options that are available for treating a given medical condition for a particular set of patients" (28, p. 625). While randomized clinical trials are the gold standard for determining treatment efficacy, observational studies may extend the results of clinical trials by examining interventions and outcomes in groups in which trials have not been or may not be performed. For example, natural experiments observing outcomes after changes in smoking policies in Pueblo, Colorado, led to the conclusion that smoking bans reduce the incidence of acute myocardial infarction (29). However, we must keep in mind that observational studies are inherently limited by selection and confounding biases, meaning that we should have the discipline to consider observational findings as hypothesis-generating, not definitive (30).

RESEARCH SUPPORT

Since the beginning of the Institute (then called the National Heart Institute) in 1948, the NHLBI has made a substantial investment in population-based cohort studies. Its current epidemiology program within the Epidemiology Branch, including all research and training programs, provided $79 million in grant support and $85 million in contract support in fiscal year 2011. The NHLBI supports a spectrum of research activities, including those driven by the Institute, those facilitated by having the infrastructure of existing studies in place, and those driven independently by investigators.

We are now facing a time of unprecedented budget constraints. The NIH budget was cut by 1% in fiscal year 2011 and is likely to be cut further in the near future. Even without nominal cuts, failure to keep up with inflation has led NIH buying power to fall to levels seen in 2000, before the doubling of the NIH budget. The NIH is now engaging the research community in a conversation about how best to support research during the budget crisis. Francis Collins, the NIH Director, recently described 4 possible levers, including "making the case" for the NIH, trimming spending across the board, evaluating and rearranging the research portfolio, and changing the ways in which NIH resources are managed. We have identified "results-based accountability" as an approach to evaluate the impact and return of research programs (31). It is clear that we will have to make increasingly difficult decisions about what and how much to fund; we look to the research community to see the challenges as an opportunity to develop new, more efficient and less expensive business models. This is especially important for epidemiology, which some scientists specifically cite as a possible source for large-scale funding cuts (32).

CONCLUSIONS

We hope our comments will spur the epidemiology community to engage in a serious and likely uncomfortable but hopefully stimulating exercise about how to address our 4 challenges: 1) how to discourage performance and overinterpretation of low-impact research or hopelessly confounded, false-positive findings; 2) how to develop novel, alternative strategies to support high-quality, large-scale research at low cost; 3) how to infuse new technologies and hypotheses within a rapidly changing environment; 4) and how to employ strategic thinking to identify the highest-priority directions. While we are aware of some approaches, we look to the research community to employ its strengths and diversity in its conversations with us.

To facilitate this discussion, we invite all readers and all of the epidemiology community to respond to the 4 challenges on a Web log that has been instituted by the NHLBI. The Web link is http://www.nhlbi.nih.gov/forum/epi/. The *Challenges to Epidemiology* blog is maintained by the NHLBI, and we are seeking a vibrant and significant discussion of these issues. The final objective of the conversation is to provide a foundation for new directions in epidemiology at the NHLBI. While no promise of research awards can be made, we will promise that the conversation will not be an empty exercise but will influence future programmatic directions.

Using a novel playbook, the "Foldit" players solved a problem in 10 days that the scientific establishment could not tackle in over 10 years. Thinking about our upcoming conversations, it should be clear to all of us that we need a new playbook in epidemiology if the method is to have any long-term hope for exploiting its powers. At the same time, it is critical for us to learn lessons of history to assure that epidemiology plays the right role in the scientific enterprise. Years ago investigators identified hypertension and hypercholesterolemia as risk factors for coronary disease, but they appropriately cautioned that their observational findings alone could not, indeed should not, direct practice (33). These investigators recognized that epidemiology lends valuable insights into disease mechanisms and prognosis but usually cannot directly inform clinical practice. While looking forward to a new era of epidemiology and mega-epidemiology, we also need to see a return to that kind of humility. As a coach might write in the playbook, "Be innovative, but remember the fundamentals."

ACKNOWLEDGMENTS

Author affiliations: Epidemiology Branch, Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, Maryland (Paul D. Sorlie); Program in Prevention and Population Sciences, Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, Maryland (Diane E. Bild); and Office of the Director, Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, Maryland (Michael S. Lauer).

The authors acknowledge the valuable input of Drs. Jean Olson and Richard Fabsitz of the Epidemiology Branch of the National Heart, Lung, and Blood Institute.

Conflict of interest: none declared.

REFERENCES

1. Khatib F, DiMaio F, Cooper S, et al. Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nat Struct Mol Biol.* 2011;18(10):1175–1177.
2. Boyle A. *Gamers Solve Molecular Puzzle That Baffled Scientists*. New York, NY: MSNBC; 2011. (http://cosmiclog.msnbc.msn.com/_news/2011/09/16/7802623-gamers-solve-molecular-puzzle-that-baffled-scientists). (Accessed January 27, 2012).
3. Kuller LH. Circular epidemiology. *Am J Epidemiol*. 1999;150(9):897–903.
4. Ioannidis JP, Tarone R, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology.* 2011;22(4):450–456.
5. Nabel EG. National Heart, Lung, and Blood Institute: transitions. *Circulation*. 2009;119(10):1453–1455.
6. National Institutes of Health. *Enhancing Peer Review at NIH*. Bethesda, MD: National Institutes of Health; 2011. (http://enhancing-peer-review.nih.gov. (Accessed December 5, 2011).
7. Chambless LE, Folsom AR, Sharrett AR, et al. Coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC) study. *J Clin Epidemiol*. 2003;56(9):880–890.
8. Ridker PM, Paynter NP, Rifai N, et al. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation*. 2008;118(22):2243–2251.

9. National Heart, Lung, and Blood Institute. *Shaping the Future of Research: A Strategic Plan for the National Heart, Lung, and Blood Institute*. Bethesda, MD: National Heart, Lung and Blood Institute; 2007. (NIH publication no. 07-6150).
10. National Heart, Lung, and Blood Institute. *Working Group Summary. The Next Step Population Studies in the "-OMIC" Age*. Bethesda, MD: National Heart, Lung, and Blood Institute; 2005. (http://www.nhlbi.nih.gov/meetings/workshops/population.htm. (Accessed December 6, 2011).
11. Benjamin EJ, Chen PS, Bild DE, et al. Prevention of atrial fibrillation: report from a National Heart, Lung, and Blood Institute workshop. *Circulation*. 2009;119(4):606–618.
12. Khoury MJ, McBride CM, Schully SD, et al. The Scientific Foundation for Personal Genomics: recommendations from a National Institutes of Health-Centers for Disease Control and Prevention multidisciplinary workshop. *Genet Med*. 2009;11(8):559–567.
13. Kaltman JR, Thompson PD, Lantos J, et al. Screening for sudden cardiac death in the young: report from a National Heart, Lung, and Blood Institute working group. *Circulation*. 2011;123(17): 1911–1918.
14. Califf RM, Ginsburg GS. Organizational improvements to enhance modern clinical epidemiology. *JAMA*. 2008;300(19): 2300–2302.
15. Yeh RW, Sidney S, Chandra M, et al. Population trends in the incidence and outcomes of acute myocardial infarction. *N Engl J Med*. 2010;362(23):2155–2165.
16. Persell SD, Wright JM, Thompson JA, et al. Assessing the validity of national quality measures for coronary artery disease using an electronic health record. *Arch Intern Med*. 2006;166(20): 2272–2277.
17. Banerjee D, Chung S, Wong EC, et al. Underdiagnosis of hypertension using electronic health records. *Am J Hypertens*. 2012;25(1):97–102.
18. Howard VJ, Cushman M, Pulley L, et al. The Reasons for Geographic and Racial Differences in Stroke Study: objectives and design. *Neuroepidemiology*. 2005;25(3):135–143.
19. Collins FS, Manolio TA. Merging and emerging cohorts: necessary but not sufficient. *Nature*. 2007;445(7125):259.
20. Willett WC, Blot WJ, Colditz GA, et al. Merging and emerging cohorts: not worth the wait. *Nature*. 2007;445(7125):257–258.
21. Gaziano JM. The evolution of population science: advent of the mega cohort. *JAMA*. 2010;304(20):2288–2289.
22. Allen NE, Beral V, Casabonne D, et al. Moderate alcohol intake and cancer incidence in women. *J Natl Cancer Inst*. 2009;101(5): 296–305.
23. Rosansky SJ, Eggers PW. Trends in the US end-stage renal disease population: 1973–1983. *Am J Kidney Dis*. 1987;9(2):91–97.
24. Krumholz HM, Herrin J, Miller LE, et al. Improvements in door-to-balloon time in the United States, 2005 to 2010. *Circulation*. 2011;124(9):1038–1045.
25. Kohane IS. Using electronic health records to drive discovery in disease genomics. *Nat Rev Genet*. 2011;12(6):417–428.
26. Kho AN, Pacheco JA, Peissig PL, et al. Electronic medical records for genetic research: results of the eMERGE consortium. *Sci Transl Med*. 2011;3(79):79rel.
27. Fröbert O, Lagerqvist B, Gudnason T, et al. Thrombus Aspiration in ST-Elevation myocardial infarction in Scandinavia (TASTE trial). A multicenter, prospective, randomized, controlled clinical registry trial based on the Swedish Angiography and Angioplasty Registry (SCAAR) platform. Study design and rationale. *Am Heart J*. 2010;160(6):1042–1048.
28. Lauer MS, Hodes R. Epidemiology, comparative effectiveness research, and the National Institutes of Health: forces for health. *Epidemiology*. 2011;22(5):625–628.
29. Bartecchi C, Alsever RN, Nevin-Woods C, et al. Reduction in the incidence of acute myocardial infarction associated with a city-wide smoking ordinance. *Circulation*. 2006;114(14):1490–1496.
30. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124. (doi:10.1371/journal.pmed.0020124).
31. Lauer MS. Thought exercises on accountability and performance measures at the National Heart, Lung, and Blood Institute (NHLBI): an invited commentary for circulation research. *Circ Res*. 2011;108(4):405–409.
32. Rosbash M. A threat to medical innovation [editorial]. *Science*. 2011;333(6039):136.
33. Lauer MS. What now with screening electrocardiography? *Ann Intern Med*. 2011;155(6):395–397.