



# HHS Public Access

Author manuscript

*Am J Epidemiol.* Author manuscript; available in PMC 2015 December 16.

Published in final edited form as:

*Am J Epidemiol.* 2015 December 15; 182(12): 977–979. doi:10.1093/aje/kwv228.

## Planning for the Future of Epidemiology in the Era of Big Data and Precision Medicine

Muin J. Khoury\*

Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Georgia

Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, Maryland

### Abstract

We live in the era of genomics and big data. Evaluating the impact on health of large-scale biological, social, and environmental data is an emerging challenge in the field of epidemiology. In the past 3 years, major discussions and plans for the future of epidemiology, including with several recommendations for actions to transform the field, have been launched by 2 institutes within the National Institutes of Health. In the present commentary, I briefly explore the themes of these recommendations and their effects on leadership, resources, cohort infrastructure, and training. Ongoing engagement within the epidemiology community is needed to determine how to shape the evolution of the field and what truly matters for changing population health. We also need to assess how to leverage existing epidemiology resources and develop new studies to improve human health. Readers are invited to examine these recommendations, consider others that might be important, and join in the conversation about the future of epidemiology.

### Keywords

big data; epidemiology; funding; genomics; precision medicine; training

---

A recent blog post (1) summarized the results of multiyear strategic efforts and discussions on planning for the future of epidemiology conducted by 2 institutes within the National Institutes of Health, the National Cancer Institute and the National Heart, Lung, and Blood Institute. Although the 2 institutes had independent processes, they both arrived at somewhat comparable and overlapping lists of broad and specific recommendations for actions that fall into thematic areas: leadership, resources, epidemiology cohorts, methods and technologies, workforce development, data and knowledge integration, and impact evaluation (Table 1). The primary impetus behind these recommendations is to transform the field and its funding strategies at a time of big data science (2) and technological developments but also resource constraints (3). Given the continued importance of epidemiology as a foundational science for public health and clinical practice, we at the *American Journal of Epidemiology* seek to encourage an open dialogue and sharing of ideas with our readers about the future of the

---

\*Correspondence to Dr. Muin J. Khoury, CDC Office of Public Health Genomics, 1600 Clifton Road, Atlanta, GA 30329 (muk1@cdc.gov).

Conflict of interest: none declared.

field. We feel that these recommendations and related questions for discussion shown in Table 1 are worthy of further exploration by the community and are applicable beyond the subject areas of cardiovascular disease and cancer to the field of epidemiology as a whole.

The planning for the future of epidemiology by 2 institutes within the National Institutes of Health coincides with the launch of the Precision Medicine Initiative (4) in 2015. A major component of this initiative is the establishment of a large longitudinal epidemiology cohort of a million or more participants in whom genetic and environmental determinants of a wide variety of human diseases can be studied (4).

In the coming years, crucial themes for moving the field of epidemiology forward will include sharing of resources, data, and metadata; evaluation of new methods and technologies to measure exposures, susceptibility, and outcomes; and identification of new ways of collecting personal (e.g., mobile health or “m-health”) and macro-level data. Shared resources such as whole genome sequencing of study participants will help in epidemiologic studies across age and disease spectra.

The use and evaluation of new methods and technologies in epidemiologic research, including new methods of data collection, require extra attention. For example, in the rapidly moving field of genomics and other “omics,” editors of this *Journal* have anticipated the need for education and information dissemination about various omics fields in the Practice of Epidemiology section (5). A few Primers on -Omic Technologies have been published in the past year (6–10), and more are under way.

More generally, we need to rethink workforce development and training of 21st century epidemiologists in data sciences, collaborative research, and more. This sentiment was echoed in a recent collaborative paper from the American College of Epidemiology on the need to retool epidemiologic competencies in the coming decades to ensure relevance of the field and enhance its ability to adapt to evolving global health needs (11).

At the heart of the transformation of epidemiology is perhaps a revised expectation of what the field can or cannot deliver even in the midst of a technological revolution. Undoubtedly, large-scale population studies tend to be expensive, and the integration of additional measurements will make them even more expensive. Vasan and Folsom commented on, among other things, the undue focus of the National Heart, Lung, and Blood Institute report on administrative efficiency and cost savings (12). In addition, the focus on the value of risk-factor epidemiology, which has been debated for years, is put into sharper focus in the era of precision medicine. As commented by Keyes and Galea, a focus on “precision” in risk factor analysis “could come at the expense of engagement with the broader causal architecture that produces population health” (13, p. 305). Integrating social and biological determinants in epidemiologic studies is easy in principle but difficult to achieve. Kuller et al. echoed some of the same sentiments and further criticized the development of “large mega cohorts without attention to specific hypotheses” (14, p. 1350). Clearly, epidemiologists have a vested interest in making the field more “consequentialist” (15) and “translational” (16). The epidemiology community will need to shape the evolution of the field and ultimately “what will truly matter most for changing population health” (13, p.

310). We also need to assess how to leverage existing epidemiology resources and develop new studies to improve human health. We invite readers of the *Journal* to take a look at some of the issues listed in Table 1, consider others that might be important, and join in the conversation about the future of epidemiology by submitting commentaries or letters for publications in the *Journal*.

## Acknowledgments

The opinions expressed in this paper are those of the author and not those of the Centers for Disease Control and Prevention.

## References

1. Khoury, MJ.; Wei, G. The future of epidemiology in the era of precision medicine: cancer, cardiovascular disease, and beyond [Web log post]<http://epi.grants.cancer.gov/blog/archive/2015/08-11.html>. Published August 11, 2015. Updated October 1, 2015. Accessed November 3, 2015
2. Khoury MJ, Ioannidis JP. Big data meets public health. *Science*. 2014; 346(6213):1054–1055. [PubMed: 25430753]
3. Lauer MS. Time for a creative transformation of epidemiology in the United States. *JAMA*. 2012; 308(17):1804–1805. [PubMed: 23117782]
4. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015; 372(9):793–795. [PubMed: 25635347]
5. Khoury MJ. A primer series on -omic technologies for the practice of epidemiology. *Am J Epidemiol*. 2014; 180(2):127–128. [PubMed: 24966223]
6. Mischak H, Critselis E, Hanash S, et al. Epidemiologic design and analysis for proteomic studies: a primer on -omic technologies. *Am J Epidemiol*. 2015; 181(9):635–647. [PubMed: 25792606]
7. Tzoulaki I, Ebbels TMD, Valdes A, et al. Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. *Am J Epidemiol*. 2014; 180(2):129–139. [PubMed: 24966222]
8. Nair VS, Pritchard CC, Tewari M, et al. Design and analysis for studying microRNAs in human disease: a primer on -omic technologies. *Am J Epidemiol*. 2014; 180(2):140–152. [PubMed: 24966218]
9. Foxman B, Martin ET. Use of the microbiome in the practice of epidemiology: a primer on -omic technologies. *Am J Epidemiol*. 2015; 182(1):1–8. [PubMed: 26025238]
10. Coughlin SS. Toward a road map for global -omics: a primer on -omic technologies. *Am J Epidemiol*. 2014; 180(12):1188–1195. [PubMed: 25409464]
11. Brownson RC, Samet JM, Chavez GF, et al. Charting a future for epidemiologic training. *Ann Epidemiol*. 2015; 25(6):458–465. [PubMed: 25976024]
12. Vasani RS, Folsom AR. Invited commentary: future of population studies—defining research priorities and processes. *Am J Epidemiol*. 2015; 181(6):369–371. [PubMed: 25743322]
13. Keyes K, Galea S. What matters most: quantifying an epidemiology of consequence. *Ann Epidemiol*. 2015; 25(5):305–311. [PubMed: 25749559]
14. Kuller LH, Bracken MB, Ogino S, et al. The role of epidemiology in the era of molecular epidemiology and genomics: summary of the 2013 *AJE*-sponsored Society of Epidemiologic Research Symposium. *Am J Epidemiol*. 2013; 178(9):1350–1354. [PubMed: 24105654]
15. Galea S. An argument for a consequentialist epidemiology. *Am J Epidemiol*. 2013; 178(8):1185–1191. [PubMed: 24022890]
16. Khoury MJ, Gwinn M, Ioannidis JP. The emergence of translational epidemiology: from scientific discovery to population health impact. *Am J Epidemiol*. 2010; 172(5):517–524. [PubMed: 20688899]

**Table 1**

Thematic Areas, Recommendations, and Selected Issues for Discussion in Planning for the Future of Epidemiology

Recommendation Themes	National Heart, Lung, and Blood Institute (3)	National Cancer Institute (4)	Selected Issues for Discussion
Scientific direction	Convene a scientific forum to anticipate the major scientific questions and methodological needs in epidemiology and population science over the next 10–20 years.	Transform the practice and extend the reach of epidemiology beyond initial discovery to include more translation, evaluation, and implementation.	How can we integrate multiple determinants of population health, including biological, social, and environmental factors? How can we encourage academic and research institutions to promote career advancement that rewards interdisciplinary and translational research?
Resource sharing	Create a dynamic compendium of large epidemiologic resources, including cohort studies, data sets from clinical trials, registries, biorepositories, and other relevant epidemiologic resources to assist the research community in identifying and accessing key existing resources and to improve the return on the investment from these studies.	Provide greater access to data, metadata, and specimens to foster collaboration, to ensure reproducibility and replication, and to accelerate translation into policies that impact population health.	How can we support the harmonization and quality of existing epidemiologic data (including cohorts and consortia) and the creation of new population study repositories? Should all new observational epidemiologic studies be registered (like randomized controlled trials)? How can we develop incentives for data sharing and for reproducibility and replication?
Maximization of the research potential of existing cohorts	Create a cohort consortium to support large-scale collaborations and provide a coordinated, interdisciplinary approach to address scientific questions, achieve economies of scale, create opportunities for collaboration, and accelerate the pace of research and the implementation of new methods.	Expand cohort studies across the lifespan and include multiple health outcomes. Maximize the output and productivity from existing cohorts and assess the need for new cohorts of etiology and outcomes, including multiple health-related outcomes and intermediate biomarkers.	How can a cohort consortium provide a collaborative platform for implementation of large scale multioutcome initiatives (such as the Precision Medicine Initiative)?
Methods and technologies	Actively engage in studies to establish the validity, reliability, and scalability of electronic tools for primary data collection.	Develop and validate reliable methods and technologies to quantify exposures and outcomes in massive scale and to assess concomitantly multiple factors in complex diseases.	How can epidemiology play a role in validating and integrating new methods and technologies?
Training and workforce development	Establish an adequate workforce to conduct population sciences “of the future”; one approach is to create multifaceted and complementary career development grants.	Train 21st century epidemiologists with an increasing emphasis on collaboration, multilevel analyses, data science, knowledge integration, and translation.	What new competencies are needed for 21st century epidemiologists to address emerging scientific questions and global health issues?
Integration of observational and interventional epidemiology	Where genuine efficiencies can be created, encourage the integration of clinical trials and epidemiologic studies.	Foster integration of observational epidemiologic studies with intervention trials.	How can we encourage more “experimentation” in large-scale epidemiologic cohort studies?
Evaluation and return on investment	Implement a competitive peer review–based model for its portfolio of large epidemiologic and population studies.	Support knowledge integration and meta research (systematic reviews, modeling, decision analysis, etc.) to identify gaps, inform funding, and integrate epidemiologic knowledge into decision making.  Develop and design rational cost-effective epidemiologic studies and resources to optimize funding, accelerate translation, and maximize health impact.	Is planning for the future of epidemiology too narrowly focused on administrative efficiency and resource saving?  Can metrics be developed for evaluating the success and impact of epidemiologic research?

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Adapted from Khoury and Wei. (1).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript