American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2017.

Vol. 186, No. 7
DOI: 10.1093/aje/kwx226
Advance Access publication:
July 15, 2017

Editorial

Editorial: Emergence of Gene-Environment Interaction Analysis in Epidemiologic Research

Muin J. Khoury*

* Correspondence to Dr. Muin J. Khoury, Office of Public Health Genomics, Centers for Disease Control and Prevention, 1600 Clifton Road, Mailstop E33, Atlanta, GA 30329 (e-mail: muk1@cdc.gov).

In this issue of the *Journal*, we publish 4 review articles (1–4) on gene-environment interaction (G×E) analysis in epidemiologic research. The papers resulted from a 2014 workshop held by the National Institute of Environmental Health Sciences and the National Cancer Institute to explore new approaches for discovery and characterization of G×E in epidemiologic research. The 4 papers provide an update on: 1) the state of the science in analytical methods of G×E (1); 2) opportunities for incorporation of biological knowledge into G×E analyses (2); 3) lessons learned from past G×E successes (3); and 4) overarching themes on current challenges and opportunities for gene-environment interaction studies of complex diseases (4). Topics include improved data analytical methods, environmental exposure assessment (5), and incorporation of functional information.

Together, these papers provide an overview of the G×E field at a time of explosive growth in the quantity and types of data that are being collected at the individual and population levels. Obviously, genome-sequencing data is a major emerging source of information on individual genetic susceptibility, but we are increasingly able to use other "-omic" data (6)—such as metabolomics, proteomics, epigenetics, and others—to measure and characterize biological processes that result from G×E. We are also increasingly able to join diverse biological data with various environmental, social, financial, geographic, and transactional data in a "big data" (7) health-impact framework.

Increasingly, all epidemiologists and all branches of epidemiology—not just genetic epidemiology—will be called upon to conduct G×E analyses in their studies. There are many reasons for studying G×E, including providing insights into disease biology, building better models of disease prediction and prognosis, and identifying subgroups of the population with much higher disease risks based on combined genetic and environmental factors. Also, in our search for genetic components of various diseases, the modifying effects of environmental risk factors are not often taken into account. Therefore, leveraging G×E may result in discovery of additional disease susceptibility loci. Ultimately, uncovering interactions should teach us about disease biology and lead to new opportunities for treatment and prevention.

As shown in these papers, in spite of the enthusiasm for and interest in G×E analysis, there are only a few genuine success stories from the study of G×E. Uncovering interactions is fraught with potential biases and methodological issues. Inherently, there is the problem of low statistical power when testing for G×E in studies designed to uncover main effects of variables. There is also the problem of the complexity of measuring environmental exposures and the difficulty in assigning temporality, especially in case-control studies. The problems of false positivity, data dredging, and selective reporting of positive interactions further complicate the field. Other problems include the limited range of genetic and/or environmental variation, scale dependence in the definition of statistical interaction, and a lack of biological data on the health impact of many genetic variants.

With the launch of the Precision Medicine Initiative (*All of Us* Research Program) (8)—in which a million participants will contribute myriad personal, genomic, environmental, and other data and be followed for health outcomes—the stage is set for numerous methodological and analytical developments in G×E research in the next decade. We hope the papers included here can provide a glimpse for our readers into the past, present, and potentially rich opportunities for the future of G×E analysis in epidemiologic research.

ACKNOWLEDGMENTS

Author affiliation: Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Georgia (Muin J. Khoury).

Conflict of interest: none declared.

REFERENCES

1. Gauderman WJ, Mukherjee B, Aschard H, et al. Update on the state of the science for analytical methods for gene-environment interactions. *Am J Epidemiol*. 2017;186(7):762–770.

2. Ritchie MD, Davis JR, Aschard H, et al. Incorporation of biological knowledge into the study of gene-environment interactions. *Am J Epidemiol*. 2017;186(7):771–777.
3. Ritz BR, Chatterjee N, Garcia-Closas M, et al. Lessons learned from past gene-environment interaction successes. *Am J Epidemiol*. 2017;186(7):778–786.
4. McAllister K, Mechanic LE, Amos C, et al. Current challenges and new opportunities for gene-environment interaction studies of complex diseases. *Am J Epidemiol*. 2017;186(7): 753–761.
5. Patel CJ, Kerr J, Thomas DC, et al. Opportunities and challenges for environmental exposure assessment in population-based studies [published online ahead of print July 14, 2017]. *Cancer Epidemiol Biomarkers Prev*. (doi:10.1016/j.cmpb.2003.08.003).
6. Khoury MJ. A primer series on -omic technologies for the practice of epidemiology. *Am J Epidemiol*. 2014;180(2):127–128.
7. Khoury MJ, Ioannidis JP. Medicine. Big data meets public health. *Science*. 2014;346(6213):1054–1055.
8. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9):793–795.