

Human Genome Epidemiology (HuGE): Translating Advances in Human Genetics into Population-based Data for Medicine and Public Health

by Muin J. Khoury M.D., Ph.D.

In this editorial, I highlight the role of epidemiology as a discipline that provides population-based information on the relationship between human genetic variation and various human diseases. I also invite authors to submit the results of epidemiologic investigations to *Genetics in Medicine* and thus to contribute to translating advances in human genetics research into medical and public health practice.

Gene Discovery and the Rapid Transition to Medical Practice

In the first issue of *Genetics in Medicine*, Dr. Richard King stated that the journal's vision is that genetics plays a central role in the practice of medicine through the prediction and prevention of disease and the promotion of public health (1). In order to realize this vision, there is an urgent need to begin making sense of the implications of advances in human genetics and the barrage of gene discoveries that hits us on a daily basis (see the results of a recent World Wide Web search [2] in Table 1).

Unfortunately, what happens after a new gene discovery is announced is often a haphazard mixture of scientific excitement, heightened public awareness and commercial interest in developing and marketing genetic tests. This is exemplified in the media coverage of the 1997 publication of an association between familial colorectal cancer in Ashkenazi Jews and the presence of a mutation in the APC gene (3). Table 2 shows the accompanying news coverage of this event from one WWW source (4). Although this study needs further replication and its implications in medical practice are far from clear, it does illustrate the mounting pressures for a rapid transition from gene discovery to integration in clinical practice, which could result in the premature development and offering of genetic tests.

To ensure a systematic translation of genetic research into clinical practice, the National Institutes of Health-Department of Energy Task Force on Genetic Testing recently released national recommendations on promoting safe and effective genetic testing in the United States, emphasizing the need for data throughout the translation process (5). Types of data on genetic tests include information on their analytic and clinical validity, clinical utility as well as data on post-marketing surveillance of the impact of genetic tests and services (5). These data need to be collected on an ongoing basis not only in high-risk families but in the general population. Herein lies the value of epidemiology.

What is Epidemiology?

Epidemiology has been defined in many ways and is often viewed as the scientific core of public health (6). One widely used definition is "the study of the distribution and determinants of health-related states or events in populations, and the application of this study to control health problems

"(7). Epidemiologists not only investigate outbreaks of disease in different populations but also conduct studies to determine risk factors for various diseases, identify sections of the population with highest risk in order to target prevention and intervention opportunities, and evaluate the effectiveness of health programs and services in improving the health of the population (8).

Over the last two decades, epidemiologic methods and approaches have been increasingly integrated with those of genetics in the discipline of genetic epidemiology, which seeks to find the role of genetic factors in disease occurrence in populations and families (9). On the other hand, a new brand of epidemiology has emerged, molecular epidemiology, that seeks to study the disease occurrence using biological markers of exposures, susceptibility and effects (10).

Most discoveries for gene variants are based on studies of high-risk families or selected groups. To translate the results of this genetic research into opportunities for treating and preventing disease and promoting health, population-based epidemiologic studies are increasingly needed to quantify the impact of gene variants on the risk of disease, death and disability and to identify and quantify the impact of modifiable risk factors that interact with gene variants. The results of such studies will help medical and public health professionals to better target medical, behavioral and environmental interventions. Epidemiologic studies are also required in the process of clinical validation of new genetic tests, monitor population utilization of genetic tests, and determine the safety and effectiveness of genetic tests and services in different populations.

For this to be accomplished, there must be collaboration among epidemiologists, clinical geneticists, laboratory scientists, and medical and public health practitioners from government, professional, academic, industry and consumer organizations. A combined genetic-epidemiologic approach is essential for a better understanding of disease etiology and development of molecular diagnostics. Data generated from such collaboration are urgently needed for the development of medical and public health policy. For example, issues are currently debated regarding population-based genetic testing for breast cancer in relation to BRCA1 (11); Alzheimer's disease in relation to the Apolipoprotein E-E4 allele (12); and iron overload in relation to the hemochromatosis gene (13). Given the paucity of population-based epidemiologic data regarding the frequency of, and disease risks and environmental interactions for many newly discovered human gene variants, there is concern that appropriate health policy on the use of genetic tests may not be possible.

Human Genome Epidemiology (HuGE)

A global effort was recently launched by the Centers for Disease Control and Prevention and its many partners, the Human Genome Epidemiology Network (HuGE Net) (14). HuGE Net seeks to promote global collaboration in the development and dissemination of peer-reviewed

epidemiologic information on human genes; to develop an updated and accessible knowledge base on the World Wide Web; and to promote the use of this knowledge base by health care providers, researchers, industry, government, and the public for making decisions involving the use of genetic tests and services for disease prevention and health promotion (14).

The term human genome epidemiology (HuGE) denotes an evolving field of inquiry that uses systematic applications of epidemiologic methods and approaches in population-based studies of the impact of human genetic variation on health and disease. HuGE can be viewed as the intersection between genetic epidemiology and molecular epidemiology (14). Whereas genetic epidemiology has traditionally focused on techniques to find disease genes using linkage and segregation analysis, and molecular epidemiology focuses on using biological markers in epidemiologic studies, HuGE will begin to address issues related to the post-gene discovery phase. As 50,000 to 100,000 genes and their variants are yet to be discovered, the impact of

such discoveries on epidemiologic research will be tremendous. As Shpilberg et al., simply put it: "The sequencing of the human genome offers the greatest opportunity for epidemiology since John Snow discovered the Broad Street pump" (15), a reminder of the contributions of epidemiology to controlling a cholera outbreak in London in the nineteenth century, long before the vibrio organism was identified.

The wide spectrum of topics addressed by investigators working on human genome epidemiology and selected examples are shown in Table 3. The spectrum ranges from population-based epidemiologic research on gene variants to evaluation of genetic tests and services. Ultimately, HuGE represents the application of clinical and molecular research to a population setting and involves the collaboration and contribution of numerous specialties.

We would like to invite researchers to view Genetics in Medicine as a forum to share the results of genetic-epidemiologic investigations on human genome variation and its relation to disease. We also invite authors to collaborate in the Human Genome Epidemiology Network.

Selected results from accepted peer-reviewed articles will become part of the cumulative HuGE Net knowledge base on the World Wide Web according to criteria to be specified (22). Authors can also provide succinct HuGE reviews according to already specified guidelines (22). Ultimately, these efforts will be crucial in translating advances in human genetics into population-based data for the practice of medicine in the new millennium.

References

1. King RA. Editorial. Genetics in Medicine 1998;2:1-2.
2. USA Today Web Site, <http://www.usatoday.com/life/health/genetics/lhgec000.htm>
3. Laken SJ, Petersen GM, Gruber SB, et al. Familial colorectal cancer in Ashkenazim due to hypermutable tract in APC. Nat Genet 1997;17:79-83.
4. Cancer Research Foundation of America Web Site. 1998; Aug 28;
5. Holtzman NA, Watson MS (eds) . Promoting Safe and Effective Genetic Testing in the United States. Final Report of the Task Force on Genetic Testing. Johns Hopkins University Press (in press). Also available on the WWW
6. Institute of Medicine. Committee for the Study of The Future of Public Health. National Academy of Sciences, Washington, DC, 1988.
7. Last JM. A Dictionary of Epidemiology, third edition. New York: Oxford University Press, 1995.
8. Terris M. The Society of Epidemiologic Research and the future of epidemiology. Am J Epidemiol 1992;136:909-915.
9. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of Genetic Epidemiology. New York, Oxford University Press, 1993.
10. Schulte, P.A., Perera, F.P. (eds). Molecular Epidemiology: Principles and Practice, Academic Press, New York, 1993

11. American Society of Human Genetics. Statement on genetic testing for breast and ovarian cancer predisposition. *Am. J. Genet.* 1994;55:i-iv.
12. American College of Medical Genetics/American Society of Human Genetics Working Group. Statement on use of Apolipoprotein E testing for Alzheimer disease. *JAMA.* 1995;274:1627-1629.
13. Burke W, Thomson E, Khoury MJ et al. Hereditary hemochromatosis: gene discovery and its implications for population-based screening. *JAMA* 1998;280:172-178.
14. Khoury MJ, Dorman JS. The human genome epidemiology network (HuGENet™). *Am J Epidemiol* 1998;148:1-3.
15. Shpilberg O, Dorman JS, Ferrell RE, et al. The next stage: molecular epidemiology. *J Clin Epidemiol* 1997;50:633-638.
16. Steinberg KK, Sanderlin KC, Ou CY, et al. DNA banking in epidemiologic studies. *Epidemiol Rev* 1997;19:156-162.
17. Shaw GM, Rozen R, Finnell RH et al. Genetic variation in infant methylene tetrahydrofolate reductase and risk of spina bifida: effect of maternal vitamin use. *Am J Epidemiol* 1998;148:30-7.
18. Malone KE, Daling JR, Thompson JD, et al. BRCA1 mutations and breast cancer in the general population: analyses of women before age 25 years and in women before age 45 years with first-degree family history. *JAMA* 1998;279:922-929.
19. Vandenbroucke JP, van der Meer FJ, Helmerhorst FM, Rosendaal FR. Factor V Leiden: should we screen oral contraceptive users and pregnant women? *BMJ* 1996;313:1127-1130.
20. Farrell PM, Kosorok MR, Laxova A, et al. Nutritional benefits of neonatal screening for cystic fibrosis. *N Engl J Med* 1997;337:963-69.
21. Centers for Disease Control and Prevention. Human Genome Epidemiology Network (HuGENet™) Web Site.

Table 1

Headlines for News Stories on Genetics Discoveries in USA Today (2)

- Alzheimer-related gene linked to boxers' brain damage
- Gene linked to inherited form of Parkinson's disease
- Migraine gene found; may help search for treatments
- Gene linked to weight control
- Hate veggies? Maybe your genes are to blame
- Found: A gene that controls place memory
- Real key to happiness may be heredity
- Gene discovery may lead to cancer treatment
- Scientists discover aging gene
- Gene found for heart disorder
- Communication gene influences personality
- Aggression gene found, researchers say
- More tests are finding genetic 'time bombs'
- Discovery may lead to test for glandular illness

- Report reveals breakthrough in Lou Gehrig's disease
- Violence linked to gene defect

This search was conducted on August 29, 1998 on the USA Today Health Genetics page.

Table 2

1997 News Story: " Genetic Defect Doubles Colon Cancer Risk" (5)

Researchers have found a new genetic defect present in one of every 17 American Jews that doubles a person's colon cancer risk. This mutation is now the most common cancer-associated gene defect identified in any ethnic population. Rarely present in non-Jews, the mutation appears to be responsible for about one in four cases of inherited colon cancer in Ashkenazi Jews--those of Eastern European ancestry who constitute more than 95 percent of America's six million Jewish people.

The good news is that scientists have developed a blood test, available for \$200, that can detect this genetic defect. The test is advisable for everyone in the Ashkenazi population, whether they have a family history of colon cancer or not. A positive blood test can alert people to the need for regular colon examinations, called colonoscopies, which can detect colon cancer in its earliest stages when it is prevented or treated through removal of growths in the colon.

Table 3

Categories and Examples of Human Genome Epidemiology Studies

1. Assess the prevalence of gene variants in different populations
Example: Using the National Health and Nutrition Examination Survey III, a nationally representative sample of the U.S. population to assess the prevalence of gene variants in population subgroups (16).
2. Assess the magnitude of disease risk associated with gene variants in different populations
Example: Using population-based case-control studies of neural tube defects to assess the etiologic role of an allelic variant in the methylene tetrahydrofolate reductase enzyme (17).
3. Assess the contribution of gene variants to the occurrence of the disease in different populations.
Example: Using population-based cancer registries to assess the contribution of BRCA1 mutations to the risk of breast cancer in the general population (18).
4. Assess the magnitude of disease risk associated with gene-gene and gene-environment interactions in different populations.
Example: Using case-control studies to assess the interaction between the factor V Leiden and the use of oral contraceptives in relation to the risk of venous thrombosis (19).
5. Assess the validity and utility of genetic tests in different populations.
Example: Using a randomized clinical trial to assess the impact of newborn screening for cystic fibrosis in Wisconsin on height and weight in the first 10 years of life (20).
6. Evaluate the determinants and impact of using genetic tests and services in different populations.
Example: Using newborn screening programs for sickle cell disease in three states to evaluate gaps in utilization in health care (e.g. penicillin prophylaxis) and the determinants of morbidity and mortality (21).