American Journal of
EPIDEMIOLOGY

Volume 148
Number 1
July 1, 1998

Copyright © 1998 by The Johns Hopkins University
School of Hygiene and Public Health

Sponsored by the Society for Epidemiologic Research

EDITORIAL

The Human Genome Epidemiology Network

Muin J. Khoury[1] and Janice S. Dorman[2]

In this editorial, we outline the need and rationale for a population-based epidemiologic approach to the human genome and describe the launching of a global collaboration, the Human Genome Epidemiology Network (HuGE Net).

THE HUMAN GENOME PROJECT AND THE NEED FOR POPULATION-BASED EPIDEMIOLOGIC RESEARCH

Progress on the Human Genome Project has led to an explosion of genetic information (1). Of the estimated 100,000 human genes, more than 9,000 have been discovered, and more than 5,000 have been mapped to specific chromosomes (2). There are approximately 500 genes for which there are genetic tests used in medical practice (3). Almost daily, discoveries are announced for gene variants that affect the risks of diseases of major public health importance, from adult chronic diseases such as cancer and diabetes mellitus to infectious and immunologic disorders, as well as diseases affecting the health of infants and children. Most discoveries of gene variants are based on studies of high-risk families or selected groups. In order 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 better target medical, behavioral, and environmental interventions.

Epidemiologic studies are also required to clinically validate new genetic tests; to measure the disease positive and negative predictive values of new tests; to monitor utilization of genetic tests; and to determine the safety and effectiveness of genetic tests and services in different populations (4). Given the massive amount of population-based data that will be generated over the next few decades, it is clear that a coordinated global effort is needed to disseminate human genome epidemiologic information and to help clinicians and researchers keep up with the progress of the Human Genome Project. For this to be accomplished, there must be collaboration among epidemiologists, clinical geneticists, basic scientists, and medical and public health practitioners from governmental, professional, academic, industrial, and consumer organizations worldwide.

Epidemiologists and geneticists must play key roles involving collaboration with government agencies, industry, professional organizations, and consumer groups. At the present time, much of the ongoing clinical genetic research is not population-based, and much of the population-based epidemiologic research does not assess genetic risk factors. A combined genetic-epidemiologic approach is essential for a better understanding of disease etiology and for the development of molecular diagnostics. Data generated from such collaboration are urgently needed for the

Received for publication February 18, 1998, and in final form April 22, 1998.

Abbreviation: HuGE Net, Human Genome Epidemiology Network.

[1] Office of Genetics and Disease Prevention, Centers for Disease Control and Prevention, Atlanta, GA.
[2] Department of Epidemiology, University of Pittsburgh School of Public Health, Pittsburgh, PA.
Reprint requests to Dr. Muin J. Khoury, Office of Genetics and Disease Prevention, Centers for Disease Control and Prevention, Mailstop K28, 4770 Buford Highway, Atlanta, GA 30341.

development of medical and public health policy. For example, current debates are focusing on issues regarding population-based genetic testing for breast cancer in relation to the *BRCA1* gene (5), Alzheimer's disease in relation to the apolipoprotein E-E4 allele (6), and iron overload in relation to the hemochromatosis gene (7). Given the paucity of population-based epidemiologic data regarding frequency, disease risks, and environmental interactions for many newly discovered human gene variants, there is concern that the formation of appropriate health policy on the use of genetic tests may not be possible.

HUMAN GENOME EPIDEMIOLOGY

In recent years, we have seen the emergence of the fields of genetic epidemiology (8) and molecular epidemiology (9). While some view the two fields as synonymous, others perceive major differences in their evolution, objectives, and focus (9). In fact, the authors of this editorial have themselves had such conflicting opinions! However, there is one issue that we both agree on: the critical role of population-based studies in translating findings of the Human Genome Project into applications for medicine and public health.

We use the term "human genome epidemiology" to denote an emerging 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. We view human genome epidemiology as the intersection between molecular epidemiology and genetic epidemiology. The spectrum of topics addressed by investigators working in this area (table 1) ranges from population-based epidemiologic research on gene variants to evaluation of genetic tests and services. The need to begin addressing the range of questions raised by this specialty was recently emphasized in a one-day workshop held at the Centers for Disease Control and Prevention (Atlanta, Georgia) in which representatives from federal, state, academic, and industrial organizations discussed and endorsed the vision of HuGE Net.

TABLE 1. Goals of human genome epidemiology

1. Assess the prevalence of gene variants in different populations.
2. Assess the magnitude of disease risks associated with gene variants in different populations (relative and absolute risks).
3. Assess the contribution of gene variants to the occurrence of the disease in different populations (attributable risks).
4. Assess the magnitude of disease risks associated with gene-gene and gene-environment interactions in different populations.
5. Assess the validity of genetic tests in different populations (disease positive and negative predictive values).
6. Evaluate the magnitude and determinants of the utilization of genetic tests and services in different populations.
7. Evaluate the impact of genetic tests and services on morbidity, disability, mortality, and cost in different populations.

HUMAN GENETIC INFORMATION ON THE WORLD WIDE WEB

Given the fast pace of scientific research in medical genetics, Sikorski and Peters (10) recently predicted that much of the up-to-date information required to practice medicine in the post-Human Genome Project era will most likely be found on the Internet. Although no single Internet site can provide comprehensive information, an amazing depth of content is already on the World Wide Web. The available online information systems include systems for obtaining gene mapping information (e.g., the Human Genome Data Base), gene sequence information (e.g., GenBank), gene nomenclature, gene mutation information (e.g., the Human Gene Mutation Database), and comprehensive phenotype and gene-specific clinical and research information (e.g., Online Mendelian Inheritance in Man (OMIM) and GENLINE). Links to online genetic databases can be found on the home page of the Centers for Disease Control and Prevention's Office of Genetics and Disease Prevention (http://www.cdc.gov/genetics/).

None of the current online knowledge bases, however, provide population-specific epidemiologic data on disease risks, allele frequencies, laboratory quality, or the safety and effectiveness of genetic tests. The proposed network is an effort that is much needed to complement basic research and clinical sources of genomic information and to fill the widening gap between genetic research findings and medical and public health practice.

THE HUMAN GENOME EPIDEMIOLOGY NETWORK

It is upon this basis that the vision for HuGE Net has emerged. HuGE Net is a logical outcome of the increasing interaction between medicine and public health, as reflected in the Centers for Disease Control and Prevention's recent strategic plan (11). HuGE Net represents the collaboration of individuals and organizations from diverse backgrounds who are committed

to the development and dissemination of population-based epidemiologic information on the human genome.

The mission of HuGE Net is threefold: 1) to establish an information exchange network that promotes global collaboration in the development and dissemination of peer-reviewed epidemiologic data on human genes; 2) to develop a continually updated and accessible knowledge base on the World Wide Web; and 3) to promote the use of this knowledge base by health care providers, researchers, industry, government, and the public for decision-making involving the use of genetic tests and services.

A website is being created at www.cdc.gov/genetics to facilitate collaboration, communication, and information exchange. The development and implementation of HuGE Net will be carried out under the direction of a steering committee and multidisciplinary working groups worldwide. An informatics group will design and create the database portion of the HuGE Net knowledge base. The website will house the cumulative (and frequently changing) information on epidemiologic aspects of human genes. Ultimately, the HuGE Net site will contain: 1) review articles supplying concise peer-reviewed information on the epidemiologic aspects of human genes (see accompanying announcement); 2) original research articles of the types listed in table 1; 3) systematic reviews on such topics as technology assessments, meta-analyses, workshop summaries, and consensus conferences; 4) updated medical literature searches for HuGE Net information on specific genes; 5) links to Internet sites and databases that provide useful population-based information on specific genes worldwide; 6) commentaries, editorials, and opinion pieces; 7) announcements of conferences, workshops, and training opportunities in human genome epidemiology; 8) funding opportunities in human genome epidemiology; and 9) a forum for communication and dialogue.

Additional information on how to join the HuGE Net collaboration can be found at the website, www.cdc.gov/genetics. Authors interested in contributing HuGE Net reviews can consult the instructions in the accompanying announcement.

ACKNOWLEDGMENTS

The authors appreciate the participation and comments of the following individuals who attended the Centers for Disease Control and Prevention HuGE Net workshop on November 20, 1997: Dr. Larry Gelbert, Bristol-Myers Squibb Company (Princeton, New Jersey); Dr. Ronald LaPorte, Department of Epidemiology, University of Pittsburgh Graduate School of Public Health (Pittsburgh, Pennsylvania); Dr. Eugene Lengerich, North Carolina Department of Health and Human Services (Raleigh, North Carolina); Dr. Michele Puryear, Maternal and Child Health Bureau, Health Resources and Services Administration (Rockville, Maryland); Dr. Daniel W. Nebert, Center for Environmental Genetics and Departments of Environmental Health and Pediatrics, University of Cincinnati (Cincinnati, Ohio); Drs. Iris Obrams and Susan Nayfield, National Cancer Institute (Bethesda, Maryland); Dr. Roberta A. Pagon, Division of Medical Genetics, University of Washington School of Medicine (Seattle, Washington); Dr. Robert J. Robbins, Vice President for Information Technology, Fred Hutchinson Cancer Research Center (Seattle, Washington); and Dr. Allen Wilcox, National Institute of Environmental Health Sciences (Research Triangle Park, North Carolina).

REFERENCES

1. Ellsworth DL, Hallman DM, Boerwinkle E. Impact of the Human Genome Project on epidemiologic research. Epidemiol Rev 1997;19:3–13.
2. Online Mendelian Inheritance in Man (OMIM). Bethesda, MD: National Center for Biotechnology Information, National Library of Medicine. (http://www.ncbi.nlm.gov/Omim/).
3. Pagon RA, Covington M, Tarczy-Hornoch P. HELIX: a directory of medical genetics laboratories. (http://www.hslib.washington.edu/helix/).
4. Khoury MJ. Genetic epidemiology and the future of disease prevention and public health. Epidemiol Rev 1997;19:175–80.
5. Burke W, Daly M, Garber J, et al. Recommendations for follow-up care of individuals with an inherited predisposition to cancer. II. *BRCA1* and *BRCA2*. JAMA 1997;277:997–1003.
6. American College of Medical Genetics/American Society of Human Genetics Working Group. Statement on use of apolipoprotein E testing for Alzheimer disease. JAMA 1995;274:1627–9.
7. McDonnell SM, Witte D. Hereditary hemochromatosis: preventing chronic effects of this underdiagnosed disorder. Postgrad Med 1997;102:83–5, 88–91, 94.
8. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of genetic epidemiology. New York, NY: Oxford University Press, 1993. (Monographs in epidemiology and biostatistics, vol 22).
9. Dorman JS, International Molecular Epidemiology Task Force. Molecular epidemiology and DNA technology transfer: a program for developing countries. In: Dorman JS, ed. Standardization of epidemiological studies of host susceptibility. New York, NY: Plenum Press, 1994:241–56.
10. Sikorski R, Peters R. Genomic medicine: Internet resources for medical genetics. JAMA 1997;278:1212–13.
11. Task Force on Genetics in Disease Prevention. Translating advances in human genetics into public health action: a strategic plan. Atlanta, GA: Centers for Disease Control and Prevention, 1997. (http://www.cdc.gov/genetics/Stratpl.html).