American Journal of Epidemiology
Vol. 156, No. 4

Printed in U.S.A.
DOI: 10.1093/aje/kwf055

Genetic Test Evaluation: Information Needs of Clinicians, Policy Makers, and the Public

Wylie Burke[1], David Atkins[2], Marta Gwinn[3], Alan Guttmacher[4], James Haddow[5], Joseph Lau[6], Glenn Palomaki[5], Nancy Press[7], C. Sue Richards[8], Louise Wideroff[9], and Georgia L. Wiesner[10]

[1] University of Washington, Seattle, WA.
[2] Agency for Healthcare Research and Quality, Bethesda, MD.
[3] Centers for Disease Control and Prevention, Atlanta, GA.
[4] National Human Genome Research Institute, Bethesda, MD.
[5] Foundation for Blood Research, Scarborough, ME.
[6] New England Medical Center, Boston, MA.
[7] Oregon Health Sciences University, Portland, OR.
[8] Baylor College of Medicine, Houston, TX.
[9] National Cancer Institute, Bethesda, MD.
[10] Case Western Reserve University, Cleveland, OH.

Received for publication September 5, 2001; accepted for publication June 13, 2002.

Growing knowledge about gene-disease associations will lead to new opportunities for genetic testing. Many experts predict that genetic testing will become increasingly important as a guide to prevention, clinical management, and drug treatment based on genetic susceptibilities. As part of a Human Genetic Epidemiology workshop convened by the Centers for Disease Control and Prevention, a group of experts evaluated the evidence needed when considering the appropriate use of new genetic tests. Because new tests are likely to vary in their predictive value, their potential to direct prevention or treatment efforts, and their personal and social consequences, the task of determining appropriate use will require careful consideration of a variety of factors, including the analytic validity, clinical validity, clinical utility, and ethical, legal, and social implications of the test. Standardized formats are needed to summarize what is known and not known about new genetic tests with respect to each of these features. Following criteria for the objective assessment of test properties, reports should be structured to enable policy makers, clinicians, and the public to identify the available evidence, so that uncertainties can be taken into account when considering test use and planning future research. *Am J Epidemiol* 2002;156:311–18

factor V; genetic markers; genetic predisposition to disease; genetic screening; genetics; phenylketonurias

Abbreviations: HNPCC, hereditary nonpolyposis colorectal cancer; HuGE, Human Genome Epidemiology.

Understanding the genetics of human health typically begins with the identification of genetic variants associated with specific diseases. Epidemiologic studies play a key role in identifying these associations. Assigning causality can be difficult, however, because of the multifactorial nature of most diseases. Even when a genetic mutation conferring increased risk is present, health outcomes may be influenced by a variety of environmental exposures, behaviors, and other genes, and interaction among some or all etiologic factors may occur. The analytic complexity posed by studies of gene-disease associations in multifactorial disease is the subject of a related paper (1). The present paper considers the information needs that arise when a genetic risk is sufficiently established that its assessment is considered in clinical or public health practice, that is, the point at which information about a gene-disease association becomes the basis for a genetic test.

The issues discussed by Little et al. (1) concerning the

Correspondence to Dr. Muin J. Khoury, Office of Genetics and Disease Prevention, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, MS-K28, Atlanta, GA 30341-3724 (e-mail: muk1@cdc.gov).

interpretation of gene-disease associations remain relevant at this transition point, because they influence the interpretation of test results. Similarly, an understanding of any modifying environmental factors that influence the clinical effect of gene variants is important. However, a new issue arises when a genetic test is considered for clinical or public health use: whether the test provides a health benefit.

Because benefit can be evaluated only in the context of specific health outcomes, the starting point for considering the use of a genetic test is a well-defined clinical problem, for which the test is expected to improve care. The test may help direct workup or management of a clinical problem, identify candidates for specific interventions, or provide diagnostic or prognostic information. In particular, genetic testing is expected to provide unparalleled new opportunities for the prevention of disease and adverse drug reactions, through the identification of people with genetic susceptibilities (2, 3).

Over the next decade a growing number of genetic tests will be marketed to clinicians as a result of expanding genetic knowledge. Many of these new tests will use DNA-based technology, but any laboratory test used primarily to identify an inherited condition is considered a genetic test (4). More than 800 genetic tests are currently available or in research development (GeneTests-GeneClinics; http://www.geneclinics.org). Although most currently available tests are for rare diseases, tests to identify inherited risk for common diseases such as breast and colorectal cancer, thromboembolism, Alzheimer's disease, and coronary heart disease have been developed.

As new genetic tests become available, policy makers, clinicians, and the public will need to make decisions about test use. New tests are likely to vary in their predictive value, their potential to direct prevention or treatment efforts, and their personal and social consequences. Therefore, the task of determining the appropriate use for a given test will require careful consideration of a variety of factors. In developing practice guidelines, different professional organizations, agencies, and health care systems will utilize a range of methodologies to evaluate evidence and reach final decisions about the indications for test use (5–7). Guidelines procedures also vary in their consideration of costs and other social issues, in part because of differences in health care priorities and resources in different health care settings. However, all practice guidelines procedures require clear and accurate information about the test under consideration. In addition, the appropriate implementation of most practice guidelines involves clinical judgment and consideration of patient preferences; thus, clinicians and patients also need accessible and accurate information about genetic test properties. To aid policy makers, clinicians, and the public in the prudent use of new genetic tests, we propose a standardized approach to reporting the data defining test properties. Our discussion of the issues to be addressed in the reporting process is illustrated by examples of genetic tests with high and low predictive value.

PHENYLKETONURIA: HIGH EXPECTATIONS FOR IMPROVED PREVENTION

Newborn screening for phenylketonuria represents a model for the use of genetic testing for disease prevention. This public health genetics program identifies newborns with phenylketonuria, a rare genetic condition in which failure to metabolize dietary phenylalanine leads to severe mental retardation. When affected newborns are provided with a phenylalanine-restricted, tyrosine-supplemented diet, mental retardation can be prevented. To a large extent, the current enthusiasm for genetic testing derives from the belief that similar benefits may occur with the use of predictive genetic tests for a wide array of other diseases.

The success of newborn screening for phenylketonuria can be attributed to four factors: a reliable and accurate screening process; a readily identifiable population for screening; an effective treatment for those who test positive; and timely initiation of treatment (8). The current phenylketonuria screening protocol, which includes follow-up testing for confirmation after an initial elevated phenylalanine level, leads to highly sensitive and specific identification of children with phenylketonuria. However, initial screening efforts did not have this level of accuracy. In the initial phases of phenylketonuria screening, some children with moderate phenylalanine elevations were classified as affected and suffered adverse consequences from a phenylalanine-restricted diet (9).

Further, although the phenylalanine-restricted diet is effective in preventing mental retardation, it is not without burdens. Children must be taught to avoid many common foods and to adhere to a diet that is not palatable to most people. Social burdens are involved; a child with phenylketonuria must be taught to avoid the hamburgers and ice cream that may be offered at a friend's house and to handle the inevitable explanations. Education and follow-up are needed to enhance compliance, with special attention to the transition to self-management that occurs in adolescence. In addition, subsidized programs for disadvantaged families are necessary to ensure that the diet is available to all newborns that test positive.

Finally, it was not until newborn screening for phenylketonuria had been instituted that an additional complication of phenylketonuria was identified: microcephaly in children of mothers affected with phenylketonuria whose phenylalanine levels were elevated during pregnancy. This outcome led to recognition of the need for tight dietary control before and during pregnancy in women with phenylketonuria. Similarly, ongoing follow-up has demonstrated that all patients with phenylketonuria benefit from continued dietary control in adulthood, illustrating the long-term social consequences of this disorder.

These experiences emphasize that data on a broad range of issues are relevant to the evaluation of a genetic test. These include not only the expected clinical measures—the reliability of the laboratory measurement, the predictive value of the test, and the effectiveness of medical interventions—but also social issues such as access to care and the personal and familial implications of testing and interventions. The phenylketonuria experience also underscores the need for

ongoing data collection after a test is introduced to identify unanticipated consequences of the testing process.

GENETIC TESTS WITH LIMITED PREDICTIVE VALUE: FACTOR V LEIDEN AND VENOUS THROMBOEMBOLISM

Can the phenylketonuria experience be replicated with other genetic tests? It is unlikely that genetic tests for common disorders will have the predictive value of phenylketonuria testing or lead to interventions that are as definitive. This is because most common disorders are multifactorial: Many different genetic variants may contribute to risk in conjunction with environmental risk factors. As a result, genetic variants associated with common diseases usually indicate increased risk rather than certainty of future disease, and the interaction of genetic and nongenetic factors is often important in determining the risk status of a particular individual. In addition, the experience to date suggests that genetic tests associated with an increased risk for common diseases may be offered for clinical use well before outcome studies are available to determine the benefits of interventions for those with genetic susceptibility (10, 11).

Factor V Leiden, the DNA-based test most commonly ordered in the United States, offers an example to illustrate these issues. This genetic trait is a variant of the factor V gene, which codes for a protein involved in clot formation. It is present in approximately 5 percent of European Americans, 2 percent of Hispanic Americans, and 1 percent of African and Asian Americans (12). Many studies, done predominantly in populations of European descent, have now documented its association with an increased risk of venous thromboembolism, often in the setting of other circumstances that affect risk, including pregnancy, use of oral contraceptives, and surgery (12–16). Estimates of the increased risk for venous thromboembolism among carriers of factor V Leiden vary from two- to eightfold (16, 17).

Factor V Leiden is one of several genetic factors that contribute to risk for venous thromboembolism (16), and interactions between genetic and nongenetic risk factors for venous thromboembolism are well established (16, 18). For example, individuals carrying both factor V Leiden and the 202210A variant in the prothrombin gene have a substantially higher risk of venous thromboembolism than do carriers of either trait alone (17, 19, 20). In more than half of a series of families with antithrombin III deficiency, another genetic risk factor, the high risk of venous thromboembolism was due to the additional presence of factor V Leiden or the prothrombin variant (21). Thus, when venous thromboembolism occurs in the presence of factor V Leiden, it often—perhaps always—represents the effect of interactions with one or more other genetic and nongenetic risk factors (16). Estimates of the annual incidence of venous thromboembolism events in people with factor V Leiden range from 0.19 percent to 0.58 percent (13, 18, 20), suggesting that cumulative risk is in the range of 10–30 percent.

Although factor V Leiden is well established as a risk factor, its implications for clinical management are far from clear. Several rationales for factor V Leiden screening have been proposed. Carriers of the variant might benefit from avoiding oral contraception (14, 22) or postmenopausal estrogen (23). Anticoagulant prophylaxis might reduce the risk of venous thromboembolism when transient risk factors, such as surgery, pregnancy, or prolonged bed rest, occur (16, 20, 22). However, these interventions all pose potential risks as well as benefits, and none has been systematically studied in people with factor V Leiden.

These various possibilities demonstrate the importance of focusing on a defined clinical problem and setting when determining the evidence needed to assess test use. Studies assessing different test purposes require different testing protocols, study populations, and outcome measures. For example, an assessment of the use of factor V Leiden testing prior to oral contraceptive use would require a population of women receiving contraceptives and would need to evaluate both contraceptive-related thrombotic risk and the outcome of alternative methods of contraception. Conversely, the assessment of testing to determine the appropriate thrombosis prophylaxis prior to surgery would rely on the measurement of postoperative outcomes in surgical patients using specified prophylaxis protocols. Each study would require appropriate control populations and measurement of other factors associated with the risk of venous thromboembolism. Without such studies, the clinical utility of testing for factor V Leiden is called into question.

DEVELOPING STANDARDIZED APPROACHES TO PRESENTING EVIDENCE

Both the phenylketonuria and factor V Leiden examples demonstrate the importance of presenting data about genetic tests in formats that allow the specific clinical outcomes of tests to be evaluated. Data to assess genetic tests for this purpose need to address three general categories of test performance (4).

Analytic validity

The term *analytic validity* refers to the accuracy with which a particular genetic characteristic (e.g., a DNA sequence variant) can be identified in a given laboratory test. One way to report analytic validity is in terms of the test's sensitivity and specificity for the genetic variant in question. Most genetic variants can be tested by a variety of protocols, and a number of technical issues arise in evaluating analytic validity. These include the assay chosen, the reliability of the assay, the degree to which reliability varies from laboratory to laboratory, and the complexity of test interpretation. Thus, an oligonucleotide probe for a single nucleotide sequence variant is a simpler test than a linkage analysis. In the latter, accuracy of the test is based on the accuracy with which samples and medical history are collected from family members, as well as on technical aspects of the testing process in the laboratory.

Often, few published data are available concerning the analytic validity of a test, and it may be difficult to judge whether different studies of the same genetic trait used comparable laboratory methods. For example, some studies of factor V Leiden utilize a DNA-based assay while others utilize a functional assay, with differences in sensitivity and specificity (22). An adequate description of the analytic

TABLE 1. Presenting data about genetic tests

Type of study	Parameters to be assessed*
All studies	Problem addressed by study
	What clinical problem is being assessed? What is the setting in which the genetic test is being used?
	Study population
	How were subjects selected? Is information provided concerning age, gender, and racial or ethnic origin? Do subjects consist of index cases only or include multiple family members?
	If a control population is included, was it selected from the same population as the cases; were matching criteria used? What inclusion and exclusion criteria were specified?
	Laboratory assay
	What was the source of samples? What variant was measured? What laboratory method was used?
Analytic validity	Reference standards
	Did the study include samples with known genotypes, with and without the variant being assayed? What was the source of reference standards? What criteria were used to define genotypes?
	Laboratory performance
	How was the reproducibility of the assay assessed? Was the reliability of the assay assessed in a routine clinical laboratory setting?
Clinical validity	Study design
	Were data collected prospectively or retrospectively? Did the study include measurement of potential modifying factors?
	Clinical and other endpoints
	What case definition was used? What endpoints were measured? Was interpretation of endpoints blinded? Were negative results verified?
Clinical utility	Intervention
	What interventions were used? What were the criteria for use of the intervention?
	Study design
	Were the data collected prospectively or retrospectively? Was an experimental study design used? If so, was a randomization method used? Was intervention blinded?
	Clinical and other endpoints
	What outcomes were measured? Was interpretation of endpoints blinded? Were negative results verified?

* See detailed discussion of methodological considerations in Little et al. (Am J Epidemiol 2002;156: 300–10).

validity of a test requires systematic collection of data for this purpose, using defined populations with known genotypes (table 1). Proficiency testing, such as that performed under the auspices of the American College of Medical Genetics and the College of American Pathologists, provides one source of objective data on analytic validity, but proficiency testing usually involves small numbers and few studies of normal genotypes. The scarcity of published studies on analytic validity is an important limitation in the available evidence about genetic tests.

Clinical validity

The term *clinical validity* describes the accuracy with which a test predicts a particular clinical outcome. When a test is used diagnostically, clinical validity measures the association of the test with the disorder. When a test is used to identify genetic susceptibility, clinical validity measures the accuracy with which it predicts a future clinical outcome. This test property corresponds to the gene-disease association measured in epidemiologic studies (1). A useful way to present clinical validity data for clinicians is in terms of the positive and negative predictive value of the test for the occurrence of disease within a defined population.

Clinical validity is often uncertain for new genetic tests. Estimates may be based on studies in limited and potentially biased populations. Uncertainty about clinical validity is an important consideration for policy makers, clinicians, and patients, particularly when they are deciding whether the risk information provided by the test provides a sufficient ratio-

nale for the use of interventions with putative but unproven benefits.

Mutations associated with disease susceptibility are often first defined in "high risk" families, characterized by multiple affected family members, and these studies may overestimate risk. The *BRCA1* and *BRCA2* genes, for example, were first identified in families selected for the presence of breast cancer in women aged less than 60 years or ovarian cancer in four or more family members (24). The lifetime risk of breast cancer is estimated to be about 85 percent in mutation carriers from these high-risk families. However, studies using less selected populations have estimated the lifetime risk associated with *BRCA1* and *BRAC2* mutations to be 36–56 percent (25–28). These observations suggest that high-risk families represent the severe end of a spectrum of risk associated with *BRCA1* and *BRAC2* mutations and that other factors, both genetic and nongenetic, may modify risk (29). Thus, variation in clinical validity may be due to complex gene-gene and gene-environment interactions, with the result that the predictive value of many genetic tests may remain imprecise even when all appropriate studies have been completed.

Important variables in evaluating evidence about clinical validity include the size and selection criteria for the study population, the type of test used (and its analytic validity), the study design, and the clinical and other endpoints measured (table 1). Careful definition of the measures used to define the clinical outcome is needed. Studies will provide more convincing results when they also include measurement of nongenetic factors that contribute to the clinical outcome. Key methodological considerations include the comparability of case and control populations (where a case-control design is used), whether the interpretation of endpoints was blinded, and whether negative results were verified.

Many genetic variants affect more than one clinical outcome; for example, factor V Leiden is associated with both venous thromboembolism and pregnancy loss (30), and some genetic variants increase risk for two or more different cancers (24). Multiple disease endpoints need to be considered when studying such variants; clinical validity may be high for one endpoint and low for another.

Clinical utility

Clinical utility refers to the likelihood that the test will lead to an improved health outcome. Measurement of this characteristic requires evaluation of outcomes associated with testing and clinical interventions. Central questions are the effectiveness of the interventions available for people who test positive and the social consequences of testing for people with positive and negative test results. Standardized approaches are available to assess the quality of evidence concerning interventions (31).

As with information about clinical validity, knowledge about optimal prevention strategies is accumulated incrementally. As is usual with emerging technology, people most genetically susceptible are initially offered interventions based on clinical reasoning or extrapolation from data in other populations, for example, early initiation of mammography screening in women with *BRCA1* and *BRAC2* mutations. The risks and benefits of these interventions can be understood only after systematic observation, in the form of well-designed controlled trials, cohort studies, or case-control studies. Because most genetic risk factors, even for common diseases, occur in a small percentage of the population (e.g., the 1–5 percent prevalence of factor V Leiden), sample sizes of genetically susceptible subjects are often small. Definitive understanding of clinical utility may be dependent on reliable methods for the pooled analysis of different studies.

Important variables in evaluating evidence about clinical utility are the size and selection criteria for the study population, the type of laboratory assay and intervention used, and the study design (table 1). The health benefit proposed for a test will determine some of the outcome measures; others are determined by hypotheses concerning the harms of the testing process. As with clinical validity studies, the comparability of the cases and controls is an important issue when a case-control design is used. If an experimental design is used, elements of the study methodology (e.g., randomization strategy) are important. The methods for measuring clinical and other endpoints should be specified, as well as whether negative results were verified.

Because genetic variants may affect multiple different clinical endpoints, a test that has high clinical utility for one endpoint may also provide extraneous information of little utility for another. For example, a person with a mutation predisposing to hereditary nonpolyposis colorectal cancer (HNPCC) faces a high lifetime risk of colorectal cancer and is likely to benefit from early initiation of regular colonoscopy as a means to reduce colorectal cancer mortality (32). However, HNPCC is also associated with increased risk for endometrial, ureteral tract, and upper gastrointestinal cancers; methods of screening for endometrial cancer are of unproven efficacy, and no preventive care is available for the other cancers associated with HNPCC (33).

ETHICAL, LEGAL, AND SOCIAL IMPLICATIONS OF GENETIC TESTS

Genetic tests have ethical, legal, and social implications. A study that is limited to medical outcomes (e.g., whether or not a particular clinical diagnosis is present or a particular outcome of treatment occurs) will be inadequate to evaluate the full implications of testing. Some social factors, such as access to testing and treatment and their cost, are critical determinants of medical outcomes. Others may determine the potential for a testing program to cause harm.

Unfortunately, the outcomes that have generated the greatest concern about genetic testing, such as insurance or employment discrimination, stigmatization, and long-term psychologic harms from testing (34–36), are difficult to study. Separate studies designed to provide adequate measurements of these outcomes may be necessary. For example, a United Kingdom survey evaluated reports of insurance discrimination among families with a wide range of genetic disorders (37). Similarly, evaluation of costs and of inequities in access to genetic services may require studies designed to examine this issue.

TABLE 2. Evidence table for a hypothetical study of clinical utility

Population	Genetic test methodology	Study design	Intervention	Endpoints measured	Results
Selection method for intervention group and controls; size	Laboratory assay used; analytic validity if known	Key features of study design	All intervention procedures	Clinical and other endpoints	Summary of main outcomes of study
Cases: sequential patients with diagnosis *X* who accept intervention *Y* (*n* = 80)	Oligonucelotide probe of mutation in gene *Z*. No formal assessment of analytic validity	Nonrandom assignment to case-control group	New radiologic screening procedure designed for early detection of cancer in patients with diagnosis *X*. Positive findings confirmed by biopsy and treated with surgery and chemotherapy	No. of positive screens and biopsies in cases	More cancer (15% vs. 7%), earlier stage of diagnosis, and more hospitalizations in cases than in controls
Controls: patients refusing intervention *Y* (*n* = 43)		Nonblinded intervention		No. of cancer diagnoses and cancer stage in cases and controls	25% of cases had a false positive screen (radiograph positive/ biopsy negative)
		Annual in-person follow-up for 5 years		Hospitalizations	Deaths the same in both groups
				Deaths	
				No measurement of quality of life	

Understanding the factors that determine interest in predictive testing is also important. The clinical utility and personal value attached to knowledge about genetic risk may vary by age or other personal characteristics. For example, the implications of a positive test for a *BRCA1* or *BRCA2* mutation differ considerably for a woman who has not yet had children compared with one who has, because oophorectomy is an important prevention option for such women. In addition, some testing decisions may be motivated by the desire to help children; a woman with cancer may be more interested in *BRCA1* and *BRAC2* testing if she has daughters who might benefit from the information.

The appropriate application of genetic testing requires information about the costs and effects of different pretest and posttest counseling procedures and about posttest behaviors. Knowledge about genetic risk does not necessarily lead to risk-reducing behaviors and could, contrariwise, produce a fatalistic attitude that reduces motivation (38). These concerns point to the importance of studying the meaning assigned to genetic information (e.g., whether test recipients or health care providers view genetic test results differently from the results of other medical tests). For example, the media often represent genetics as deterministic (39), but whether this misunderstanding affects the interpretation of genetic test results is not known. Some experimental study designs can incorporate these empirical questions, but long-term follow-up may be needed to ensure an adequate understanding of the personal and social outcomes of the testing process.

GENETIC TEST REPORTS: MAKING SENSE OF THE EVIDENCE

Accurate reporting of the properties of genetic tests, using standardized methodologies, serves an important clinical purpose. Reports summarizing what is known and not known about a genetic test allow policy makers, health care providers, and patients to evaluate new testing opportunities with the assurance that their decisions about test use will be based on the available evidence. In-depth summaries of evidence can also guide future research.

The first step in this process is to encourage authors to present methods and results in standardized ways, as outlined in table 1 and in Little et al. (1). Brief summaries following a succinct and systematic format similar in concept to a structured abstract would provide information about the current status of genetic tests. Because knowledge about tests will change over time, these status reports would ideally be dated and revised on a periodic basis.

Two other more detailed types of reports can provide the necessary background for the brief summaries and offer additional support to policy makers, clinicians, and researchers. One is the structured review, providing a systematic evaluation of evidence on the analytic validity, clinical validity, and clinical utility of a test, including consideration of the ethical, legal, and social implications of testing. The second is the construction of evidence tables that provide ready access to information about the individual studies on which summaries or structured reviews are based. An example of an evidence table is shown in table 2. A reporting format of this kind clarifies both the strengths and the weaknesses of available data, allows studies to be

compared, and makes gaps in knowledge apparent. In the hypothetical screening study shown in table 2, for example, the evidence table makes the weaknesses of the study apparent (e.g., intervention/control status determined by patient preference; lack of assessment of analytic validity or quality of life; lack of confirmation of negative screening results) and suggests that screening in this example increases hospitalizations but does not reduce mortality over a 5-year follow-up period. In addition to providing a realistic picture of what is known about a genetic test, an evidence table identifies additional research needs and facilitates the selection of studies for pooled analyses. This format can also be readily updated by the addition of new studies to the table.

The "Screening Brief" featured in the *Journal of Medical Screening* represents an example of a brief summary format that could be applied to genetic tests, and the Human Genome Epidemiology (HuGE) reviews that are now a regular feature in the *American Journal of Epidemiology* represent an example of a structured review. Evidence tables may be used in the development of these reports but are not generally published or otherwise made available. The Internet may provide a mechanism for dissemination of this reporting format, perhaps through Web sites already providing genetic information, such as HuGE Net (http://www.cdc.gov/genomics/hugenet/default.htm), the CancerNet Physician Data Query (PDQ; National Cancer Institute, Bethesda, Maryland) summaries on cancer genetics (http://www.cancer.gov/cancer_information/doc.aspx?viewid=58D577DE-5B0B-4757-9CE6-7F898204EEC5), and GeneTests-GeneClinics reviews (http://www.geneclinics.org). Professional organizations could play a role in the development of all these information resources. Evidence tables, in particular, provide support for the development of practice guidelines.

These strategies for reporting information about genetic tests should not be viewed as replacements for clinical practice guidelines. The latter address the use of a genetic test in a particular health care setting and must take into account the value of testing for the population being served, its cost and acceptability in this setting, and its priority relative to other health care services. Tests suitable in some settings may not be suitable in others. Although procedures beyond the test reporting outlined here are needed to address these issues, objective information concerning the current state of knowledge about a test is an essential component of this process.

SUMMARY

New techniques to organize and disseminate evidence about genetic tests are an important step in support of the rational and prudent use of this new technology. To be effective, evaluation methods must be as complete and unbiased as possible. Clinicians, guidelines panels, and other policy makers must consider genetic tests in light of alternative approaches to care; that is, the health benefits and other outcomes of a genetic test must be evaluated compared with the outcomes that could be obtained in the absence of genetic testing. To make this comparison, adequate evidence concerning the analytic validity, clinical validity, and clinical utility of genetic tests is needed, including appropriate control populations and, ideally, measurement of social outcomes.

This ambitious research agenda is unlikely to be achieved rapidly or completely for most genetic tests. Even when all appropriate research is completed, the predictive value of many genetic tests may remain imprecise, and the clinical utility may be variable. As a result, methods to disseminate information about genetic tests represent an important strategy for ensuring appropriate test use. Following criteria for the objective assessment of test properties, reports should be structured to enable policy makers, clinicians, and the public to identify quickly what is known and not known about analytic validity, clinical validity, and clinical utility, so that uncertainties can be taken into account when genetic tests are considered for use.

ACKNOWLEDGMENTS

This paper represents consensus from the Centers for Disease Control and Prevention-National Institutes of Health Human Genome Epidemiology Workshop, January 2001.

REFERENCES

1. Little J, Bradley L, Bray MS, et al. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol 2002;156:300–10.
2. Collins FS. Shattuck lecture—medical and societal consequences of the Human Genome Project. N Engl J Med 1999; 341:28–37.
3. Roses AD. Pharmacogenetics and the practice of medicine. Nature 2000;405:857–65.
4. Holtzman NA, Watson MS. Promoting safe and effective genetic testing in the United States. Final report of the Task Force on Genetic Testing. J Child Fam Nurs 1999;2:388–90.
5. Institute of Medicine CoCPG. Guidelines for clinical practice: from development to use. Washington, DC: National Academy Press, 1992.
6. Woolf SH. Evidence-based medicine and practice guidelines: an overview. Cancer Control 2000;7:362–7.
7. Miller J, Petrie J. Development of practice guidelines. Lancet 2000;355:82–3.
8. Pass K. Lessons learned from newborn screening for phenylketonuria. In: Khoury MJ, Burke W, Thomson E, eds. Genetics and public health. New York, NY: Oxford University Press, 2000:385–404.
9. Paul D. The history of newborn phenylketonuria screening in the United States. In: Holtzman NA, Watson MS, eds. Promoting safe and effective use of genetic testing in the United States: final report of the Task Force on Genetic Testing. Baltimore, MD: Johns Hopkins University Press, 1998:137–60.
10. Laken SJ, Petersen GM, Gruber SB, et al. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. Nat Genet 1997;17:79–83.
11. 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*. Cancer Genetics Studies Consortium. JAMA 1997;277:997–1003.
12. Ridker PM, Miletich JP, Hennekens CH, et al. Ethnic distribution of factor V Leiden in 4047 men and women. Implications

for venous thromboembolism screening. JAMA 1997;277: 1305–7.
13. Middeldorp S, Henkens CM, Koopman MM, et al. The incidence of venous thromboembolism in family members of patients with factor V Leiden mutation and venous thrombosis. Ann Intern Med 1998;128:15–20.
14. Bloemenkamp KW, Rosendaal FR, Helmerhorst FM, et al. Enhancement by factor V Leiden mutation of risk of deep-vein thrombosis associated with oral contraceptives containing a third-generation progestagen. Lancet 1995;346:1593–6.
15. Vandenbroucke JP, Koster T, Briet E, et al. Increased risk of venous thrombosis in oral-contraceptive users who are carriers of factor V Leiden mutation. Lancet 1994;344:1453–7.
16. Rosendaal FR. Venous thrombosis: a multicausal disease. Lancet 1999;353:1167–73.
17. Meyer G, Emmerich J, Helley D, et al. Factors V Leiden and II 20210A in patients with symptomatic pulmonary embolism and deep vein thrombosis. Am J Med 2001;110:12–15.
18. Simioni P, Sanson BJ, Prandoni P, et al. Incidence of venous thromboembolism in families with inherited thrombophilia. Thromb Haemost 1999;81:198–202.
19. De Stefano V, Martinelli I, Mannucci PM, et al. The risk of recurrent deep venous thrombosis among heterozygous carriers of both factor V Leiden and the G20210A prothrombin mutation. N Engl J Med 1999;341:801–6.
20. Martinelli I, Bucciarelli P, Margaglione M, et al. The risk of venous thromboembolism in family members with mutations in the genes of factor V or prothrombin or both. Br J Haematol 2000;111:1223–9.
21. van Boven HH, Vandenbroucke JP, Briet E, et al. Gene-gene and gene-environment interactions determine risk of thrombosis in families with inherited antithrombin deficiency. Blood 1999;94:2590–4.
22. Grody WW, Griffin JH, Taylor AK, et al. American College of Medical Genetics consensus statement on factor V Leiden mutation testing. Genet Med 2001;3:139–48.
23. Rosendaal FR. Factor V Leiden and hormone replacement therapy. Presented at the CDC Factor V Leiden Workshop, Atlanta, Georgia, April 2, 2001.
24. Ford D, Easton DF, Bishop DT, et al. Risks of cancer in *BRCA1*-mutation carriers. Breast Cancer Linkage Consortium. Lancet 1994;343:692–5.
25. Fodor FH, Weston A, Bleiweiss IJ, et al. Frequency and carrier risk associated with common *BRCA1* and *BRCA2* mutations in Ashkenazi Jewish breast cancer patients. Am J Hum Genet 1998;63:45–51.
26. Struewing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of *BRCA1* and *BRCA2* among Ashkenazi Jews. N Engl J Med 1997;336:1401–8.
27. Thorlacius S, Struewing JP, Hartge P, et al. Population-based study of risk of breast cancer in carriers of *BRCA2* mutation. Lancet 1998;352:1337–9.
28. Hopper JL, Southey MC, Dite GS, et al. Population-based estimate of the average age-specific cumulative risk of breast cancer for a defined set of protein-truncating mutations in *BRCA1* and *BRCA2*. Australian Breast Cancer Family Study. Cancer Epidemiol Biomarkers Prev 1999;8:741–7.
29. Burke W, Press NA, Pinsky LE. *BRCA1* and *BRCA2*: a small part of the puzzle. J Natl Cancer Inst 1999;91:904–5.
30. Kupferminc MJ, Eldor A, Steinman N, et al. Increased frequency of genetic thrombophilia in women with complications of pregnancy. N Engl J Med 1999;340:9–13.
31. US Preventive Services Task Force. Guide to clinical preventive services. Baltimore, MD: Williams & Wilkins, 1996.
32. Renkonen-Sinisalo L, Aarnio M, Mecklin JP, et al. Surveillance improves survival of colorectal cancer in patients with hereditary nonpolyposis colorectal cancer. Cancer Detect Prev 2000; 24:137–42.
33. Burke W, Petersen G, Lynch P, et al. Recommendations for follow-up care of individuals with an inherited predisposition to cancer. I. Hereditary nonpolyposis colon cancer. Cancer Genetics Studies Consortium. JAMA 1997;277:915–19.
34. Rothenberg KH. Genetic information and health insurance: state legislative approaches. J Law Med Ethics 1995;23:312–19.
35. Billings PR, Kohn MA, de Cuevas M, et al. Discrimination as a consequence of genetic testing. Am J Hum Genet 1992;50:476–82.
36. Lapham EV, Kozma C, Weiss JO. Genetic discrimination: perspectives of consumers. Science 1996;274:621–4.
37. Low L, King S, Wilkie T. Genetic discrimination in life insurance: empirical evidence from a cross sectional survey of genetic support groups in the United Kingdom. BMJ 1998;317: 1632–5.
38. Marteau TM, Lerman C. Genetic risk and behavioural change. BMJ 2001;322:1056–9.
39. Khoury MJ, Thrasher JF, Burke W, et al. Challenges in communicating genetics: a public health approach. Genet Med 2000;2: 198–202.