Outcomes of interest in evidence-based evaluations of genetic tests

Jeffrey R. Botkin, MD, MPH[1], Steven M. Teutsch, MD, MPH[2], Celia I. Kaye, MD, PhD[3],
Maxine Hayes, MD, MPH[4], James E. Haddow, MD[5], Linda A. Bradley, PhD[5,6],
Kathleen Szegda, MS, MPH[7], and W. David Dotson, PhD[6], on behalf of the EGAPP Working Group

Abstract: Genetic tests are increasingly available for use in traditional clinical practice settings and through direct-to-consumer marketing. The need for evidence-based information and guidance on their appropriate use has never been more apparent. The independent Working Group of the Evaluation of Genomic Applications in Practice and Prevention Initiative commissions evidence-based reviews and develops recommendations to inform decision making surrounding the implementation of genetic tests and other applications of genomic technologies into clinical practice. A critical component of this analysis involves the identification and appropriate weighting of relevant health outcomes from genetic testing. Impacts of testing on morbidity and mortality are central considerations although research to document such outcomes can be challenging to conduct. In considering the broader impacts of genetic tests on the individual, familial and societal levels, psychosocial outcomes often take on increasing importance, and their systematic evaluation is a challenge for traditional methods of evidence-based review. Incorporating these types of outcomes in evidence-based processes is possible, however, and necessary to extract balanced and complete (or as complete as available data will allow) information on potential benefits and on potential harms. The framework used by the Evaluation of Genomic Applications in Practice and Prevention Working Group in considering, categorizing, and weighting health-related outcomes as applied to genomic technologies is presented here. *Genet Med* 2010:12(4):228–235.

Key Words: *genetic test, outcomes, analytic validity, clinical validity, clinical utility, molecular diagnostic techniques, screening, penetrance, expressivity, genetic counseling*

Research in genetics has been a national priority in recent decades, resulting in an increasingly detailed understanding of genome structure and function. This knowledge is expected eventually to yield a wide array of effective interventions to treat or prevent genetic conditions.[1] However, in the shorter term, it has led to an explosion in the number of new genetic tests, which are being applied in a variety of contexts, including clinical diagnosis and prognosis, population screening, health risk assessments, prenatal diagnosis and screening, tissue analysis for acquired mutations, pharmacogenomic testing, toxicogenomic testing, forensic testing, and heritage testing. Availability of this new testing has generated excitement over prospects for applications that might achieve more personalized approaches to health care,[2] but this excitement is tempered by debate on whether present technologies, and our knowledge of their utility, are sufficient to realize this goal. A hallmark of all tests is that they produce information only and we seek to adopt tests for which the beneficial impacts of the information outweigh the harmful ones. More specifically, a test is judged to be valuable when its information leads to actions or behaviors that reduce morbidity or mortality from the condition identified or when the information has some other intrinsic value. For example, a test may have value in the context of ending a diagnostic odyssey, informing life planning, or avoiding unnecessary interventions, even when no changes in treatment or preventive strategies are warranted based on test results.

As of 2009, the GeneTests website lists >1700 diseases or conditions for which genetic testing is available.[3] Although most of these tests pertain to the diagnosis of rare genetic disorders, an increasing number of tests have applications that could affect a broad population (e.g., carrier identification, testing to predict the risk of common diseases, and pharmacogenomic testing).[4] Many of these genetic tests are being actively used by physicians in clinical practice, and some are increasingly being marketed and offered directly to consumers.[5] However, what do we really know about the validity and utility of these genetic tests? What are our expectations of a genetic test and how should we determine whether those expectations are being met? Unlike the case with new drugs, biological and medical devices, the US Food and Drug Administration (FDA) does not actively regulate most genetic tests.[6,7] A small subset of genetic tests, including those marketed as kits, and components of tests called "analyte specific reagents," fall within FDA's regulatory domain.

When it does evaluate genetic tests, the FDA's primary focus is on analytic validity and, to some extent, clinical validity. Demonstration of clinical utility is not required for clearance or approval. Certification of laboratory quality occurs through Clinical Laboratory Improvement Amendments certification, state regulation, and proficiency testing of laboratories. More detailed information on these oversight groups can be found elsewhere.[8] The net consequence of the current regulatory environment is that applications of genomic technology often move quickly from the "bench" to the "bedside." The relative lack of regulatory oversight often leaves clinicians without independent assurances that genetic tests will perform as antic-

ipated (are analytically valid) and have established clinical validity and clinical utility. Gaps in the regulatory environment surrounding genetic tests correspond to areas where important gaps in knowledge regarding these tests exist, and independent incentives are needed to promote research in these areas.[8] These knowledge gaps also leave third-party payers without an evidence-base on which to make informed coverage decisions.

To address gaps in knowledge and guidance, the Office of Public Health Genomics at the Centers for Disease Control and Prevention launched a nonregulatory initiative in 2004 called Evaluation of Genomic Applications in Practice and Prevention (EGAPP).[9,10] The purpose of EGAPP is to conduct evidence-based evaluations of genetic tests and other applications of genomic technology that are available or emerging into clinical and public health practice. The primary focus of this initiative is the independent EGAPP Working Group (EWG) established in April 2005. The EWG is currently composed of 16 multidisciplinary experts in areas that include evidence-based review, genomics, clinical and laboratory practice, public health, economics, ethics, policy, and assessment of health technologies.[11] The EWG, with the support of Centers for Disease Control and Prevention-based staff from the Office of Public Health Genomics, commissions evidence-based reviews on genetic and genomic tests and develops recommendations with clear linkage to the available scientific evidence.

An important goal of EGAPP has been to draw on the successes and lessons learned through similar processes and avoid duplication of efforts. Because existing processes were already in place to address rare genetic conditions and single-gene disorders (e.g., the Collaboration Education and Test Translation Program),[12] it was decided that the scope of EGAPP would focus on disorders with the potential for a broad, population impact.[13] Initial tasks for the EWG included defining methodologies and outcomes of interest in their assessment of genetic tests. Successful methods and processes from numerous evidence-based evaluation initiatives (e.g., US Preventive Services Task Force, Cochrane Collaboration, and international health technology groups)[14–16] have been drawn on throughout the EGAPP initiative. However, in general, the relative weighting of health-related outcomes most amenable to evidence-based evaluation (e.g., "hard," clinical outcomes) is not proportional in the assessment of genetic technologies, where other factors (e.g., "soft," behavioral outcomes) often take on increased significance. Therefore, an Outcomes Subcommittee of the EWG was convened to advise the EWG on how to define outcomes to be considered in evidence reviews and to assess them in a systematic way.

This article describes the EWG conclusions with respect to test outcomes and discusses the outcomes that the EWG assesses in framing systematic reviews and in making evidence-based recommendations. Although rare and single-gene disorders currently lie beyond the routine scope of EGAPP, they were included in the Outcomes Subcommittee's deliberations to be comprehensive, and we have drawn on selected examples of rare disease testing in this report.

ASSESSMENT OF GENETIC TESTS

EGAPP has adopted a rigorous approach to its evidence reviews and the development of its recommendations.[17] Central to EGAPPs methods are the selection process for tests to be evaluated, development of an analytic framework, and the evidence collection and assessment. The components of evaluation are the ACCE criteria, standing for Analytic Validity, Clinical Validity, Clinical Utility, and Ethical, Legal, and Social

Implications (ELSI).[18,19] Each of these four characteristics of a test is critical to its overall assessment as a safe and effective clinical tool. Typically represented in the form of a wheel centered on the specific disorder and setting in which testing is used, the ACCE approach uses a set of 44+ targeted questions to guide evaluation through each component of evaluation. Analytic validity in the context of genetic testing refers to the ability of the test to correctly characterize a known target DNA sequence, structural variation, or genotype. Although important to the analysis of a genetic test, analytic aspects of tests are not considered in our assessment of outcomes. The clinical validity of tests is related to our outcomes of interest in that false-positive and false-negative results may have adverse outcomes that will impact the overall assessment of risks and benefits of testing. The outcomes of interest for EGAPPs purposes relate primarily to clinical utility and ELSI issues, namely the health consequences (benefits and harms) of responses by patients and/or family members, physicians, and society to test information.

WHAT OUTCOMES DOES EGAPP CONSIDER?

Similar to the US Preventive Services Task Force, the EWG assesses outcomes in their evaluation of the magnitude of net health benefit (benefits less harms) of testing.[20,21] This report provides a lexicon of outcomes to be applied selectively by the EWG in specific evidence reviews and in the accompanying recommendations. It is stipulated that outcomes be meaningful to patients, their families, or communities. The selection of outcomes used for an individual EGAPP review is guided by an analytic framework, which is designed to address the specific clinical scenario and describe the outcomes of concern. Herein, we review the range of outcomes that EGAPP may consider when conducting each review.

To organize the various outcomes that may result from testing, EGAPP adopted a variation of the model proposed by Tatsioni et al.[22] The outcomes of interest for a test are organized into four groups: diagnostic and prognostic thinking, therapeutic choice, patient impact, and familial and societal impact (Table 1). ACCE-targeted questions can be applied to the individual domains of the model to identify, collect, and organize relevant data on outcomes. These domains represent a sequential flow of the test result from its initial impact on the knowledge and attitudes of the patient and clinician, through the subsequent impact on health, to the eventual impact on society more broadly. Examples of health-related outcomes within individual domains are provided, and these would be ranked in the context of the selected clinical scenario for the test under review. Although some of these outcomes will be much easier to measure than others, the goal of ranking will focus on the outcome's importance in clinical decision making rather than measurability. In cases where particular outcomes are found to be challenging to measure, and yet the implications are highly relevant to clinical decision making, the gap in knowledge will be highlighted.

As an example, Table 2 shows an outcomes matrix created by the EWG for framing the evidence review and making its recommendation on CYP450 testing among adults with non-psychotic depression treated with selective serotonin reuptake inhibitors.[23,24] Ranking of outcomes by relative importance in this matrix involves some level of subjectivity. However, it provides an intuitive approach toward systematically categorizing and assessing outcomes. The specific content and rankings in this example are relevant to the specific topic under evaluation and are not intended to reflect a general prioritization that would apply to the evaluation of all genetic tests. The matrix allows visualization of the distribution of outcomes between

Table 1 Examples of types of health-related outcomes

Potential outcomes	Examples
Diagnostic thinking/health information impact	Ending diagnostic odyssey
	Knowledge of prognosis/disease course
	Long-term planning
	Distress (increased or decreased)
	Satisfaction with testing services
	Increased/decreased sense of control
	Stigmatization or discrimination
	Incidental information (unwanted information)
	Changes in family dynamics
	Cultural, ethnic identity
Therapeutic choice	Changes in preventive or therapeutic strategies
	Adherence to therapeutic regimen
	Satisfaction with treatment choice
	Health behavior (test recipients)
Patient outcome impact	Mortality
	Morbidity
	Change in response to therapy
	Incidence of adverse outcome(s) after testing
	Severity of adverse outcome(s) after testing
	Health-related quality of life
	Pregnancy termination decisions
	Prenatal interventions
Familial and societal impact	Impact on health disparities
	Health care utilization by family members
	Disabilities perspective
	Fostering genetic determinism in society
	Eugenics attitudes in society
	Technology innovation
	Population health interventions

individual, family, and society levels, and identifies areas of increased importance to patients and physicians under the clinical scenario under evaluation. The following sections provide a brief overview of outcomes of interest to the EWG in each of the domains of the outcomes model.

DIAGNOSTIC AND PROGNOSTIC THINKING

Information provided by a genetic test initially affects the clinician's and the patient's thinking about disease likelihood,

either for the symptomatic patient or for the patient concerned about future disease. An accurate diagnosis or risk assessment can guide treatments or preventive measures in many but not all instances (as will be described later). Furthermore, a specific understanding of the genes or mutations involved with the clinical condition may refine prognostic information and/or influence treatment strategies. Where effective treatment or preventive measures are available, the harms and benefits associated with those interventions are generally the primary outcomes of interest. However, even in the absence of those options, diagnostic information can potentially bring substantial psychological benefits to patients, family members, and care providers. Given the rare nature of many genetic conditions, individuals often follow a burdensome and expensive "diagnostic odyssey" before obtaining a definitive explanation.[25,26] Providing a diagnosis ends the odyssey and reduces the uncertainty. Diagnostic information also provides a prognosis, enabling life planning even in the absence of effective interventions. These outcomes may be assessed through studies of knowledge and understanding of test results by clinicians and patients and by measures of satisfaction with obtaining genetic information. Any actual behavioral responses to test information are assessed separately, as outlined below.

Knowledge per se of genetic information also may have negative outcomes. Although not unique to genetic testing, psychological distress is common, at least over the short term, for those diagnosed with a serious condition or found to be at high risk for future disease. Therefore pre- and posttesting to quantify the level of distress for those who test positive (for either a target condition or a genetic variant conferring elevated risk for a condition) may be important to measure when evaluating genetic testing for serious conditions. Understanding distress levels also is pertinent to individuals who test negative and for those who have ambiguous results. In the case of Huntington disease, research has revealed that more stress may be caused by ambiguous results than by positive results.[27] Varied emotional responses to inconclusive *BRCA1/2* results have been observed among cancer patients, and although stress levels associated with this testing have been described as generally not reaching clinical significance, more research is needed on the psychological impact of inconclusive results.[28–30] Understanding the psychological reactions to tests is also important in guiding counseling interventions. Furthermore, the psychological state of patients and clinicians may influence subsequent behaviors with respect to adherence to treatment recommendations or early detection measures. There is a substantial literature on the psychological impacts of genetic testing for adult onset conditions, including *BRCA1/2*, Lynch syndrome (also known as hereditary nonpolyposis colorectal cancer), and Huntington disease.[28,31–38] Examples such as these, involving highly penetrant mutations, may offer limited insights into the potential psychological outcomes of tests involving variants of lower penetrance, particularly those that are provided without professional oversight or formal counseling.

A second important adverse outcome in diagnostic thinking is a consequence of misinterpretation of test results leading to false reassurance or concern, needless additional testing, or inappropriate therapy. Clinicians and patients may misunderstand results, and it is possible that this may occur more frequently with genetic tests compared with nongenetic tests, due to the complexity of the information, lack of familiarity with new tests, infrequent use by individual clinicians, confusing laboratory report forms, and lack of basic genetic knowledge for both clinicians and patients.[39] In some circumstances, there even may be confusion over whether a result is positive or

Table 2 Example outcomes matrix—testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with SSRIs

Outcomes	Rank	Outcome type		Outcome relevance		
		Patient outcome	Physician outcome	Individual level	Family level	Society level
Diagnostic thinking						
Knowledge of CYP	2		●	●		
Knowledge of risk	1	●	●	●		
Ability to interpret	1		●	●		
Therapeutic choice						
Selective use of SSRI	1		●	●		
Tailored dosage	1		●	●		
Time to dosage	2		●	●		
Adherence	2	●		●		
Changes in other drugs	3		●	●		
Patient outcomes						
Morbidity						
Short-term response	1	●		●		
Long-term response	1	●		●		
Incidence of AEs	1	●		●		
Severity of AEs	1	●		●		
Other med changes	3	●	●	●		
Mortality						
Incidence of suicide	1	●		●		
Familial and societal impact						
Economic evaluation	2					●
Reimbursement/access	2			●		●
Health disparities	3					●
Relationship to disparity	3					●

SSRI, selective serotonin reuptake inhibitor; CYP, cytochrom P450; AEs, adverse events; Rank, the relative rank of this outcome on a scale of 1 to 3 (most to least important); patient outcome, outcome is primarily relevant to the patient; physician outcome, outcome is primarily relevant to the physician; individual level, outcome is relevant on an individual level; family level, outcome is relevant at the family level; society level, outcome is relevant at a societal level.

negative.[40] In other situations, genetic risk information may be over- or underinterpreted. This can be a problem for individuals with either positive or negative predictive genetic test results. As a hypothetical example, consider the potential implications for an overweight individual who learns that he harbors certain alleles that may be associated with an elevated risk of developing type 2 diabetes. In this context, recent findings suggest little to no added value in risk prediction based on inclusion of genetic risk factors along with traditional risk factors.[41–43] Could overinterpretation of genetic test results, perhaps based on a false assumption of genetic determinism, lead to a fatalistic attitude, and thus impede behavioral changes such as weight loss, aimed at preventing diabetes? Could a similar individual found not to be at elevated genetic risk for type 2 diabetes overinterpret the result as meaning that he has a negligible chance of developing late-onset diabetes and therefore opt to

maintain a sedentary lifestyle? Predictive genetic tests often are not definitive in terms of whether the individual will be affected (penetrance), at what point in the future, or with what degree of severity (expressivity). This offers an opportunity for misunderstanding the implications of test results among some patients and perhaps clinicians.

Genetic factors typically work in concert with other biological and environmental factors that lead to illness. Nevertheless, some patients may feel doomed by their results when, in fact, disease may never occur or may occur decades in the future. Conversely, some who receive negative results on a predictive test might falsely assume they are risk free for the condition when, in fact, they are likely to be at the average population risk. Those who are falsely reassured by good news on a genetic test may forego subsequent screening or risk reduction measures that remain appropriate for them. False-positive and false-

negative results result in informational errors that may lead to inappropriate psychological reassurance or concern and potentially inappropriate clinical management.

Results of unknown clinical significance are a particular problem in DNA sequencing-based genetic tests. An unusual sequence variation in an individual may represent a deleterious mutation or a benign variant.[44] These results will be reported to clinicians and patients as variants of unknown significance, and it may take months or years for sufficient knowledge to be accumulated for the laboratory to make a final determination. Ambiguous results could potentially cause substantial distress in patients, and how clinicians manage these results should be an important outcome to assess.[45]

Genetic test results also can affect how individuals think about the control they have over their lives and how they relate to family members. An important characteristic of genetic tests is their ability to provide information about the "blood" relatives of the individual tested.[46] Guilt for passing a heritable condition to the next generation is common. There also may be an element of "survivor guilt" in unaffected siblings when other siblings are affected or receive adverse test results. In considering tests for common, multigene disorders (e.g., predictive testing for the risk of cardiovascular disease), issues such as survivor guilt may be less of a concern. However, this is not always the case. Similarly, in many cases, family members may be much less involved (or not involved at all) in the testing process for pharmacogenetic and low penetrance variants, when compared with testing for highly penetrant mutations such as in Huntington disease. In other examples, such as Lynch syndrome (described in the Familial and Societal Impact section), involvement of family members in the testing process can be a critical factor in using test results to their maximal benefit toward disease prevention. Ultimately, the highly varied nature of both the disorders for which there are genetic tests, and of the tests themselves, makes it difficult to impossible to draw conclusions that will apply to a wide variety of clinical scenarios.

Those found at risk of future disease also may have a heightened sense of vulnerability with respect to social stigma or discrimination. A common reason to decline predictive genetic testing may be the fear of genetic discrimination,[47–49] although the actual number of instances of insurance or employment discrimination appears to be small.[50] The recently passed Genetic Information Nondiscrimination Act[51] may reduce both the perception and practice of discrimination. Many individuals at genetic risk for future illness choose not to undergo testing. A complete assessment of a genetic test would include an understanding of whether at-risk individuals want testing and why they make certain decisions about testing. Genetic conditions also are frequently associated with specific racial or ethnic groups. Genetic testing or screening may be targeted to subgroups in the population, and results may foster a sense of group identity and/or group vulnerability. These issues not only flow into the ELSI concerns but also are part of how people think about the results of testing.

THERAPEUTIC CHOICE

How do genetic test results impact clinician and patient behavior? If there is no response to the test information that leads to changes in the preventive or therapeutic strategies resulting in a better health outcome, then there is no clinical utility. One such circumstance arises when a strong family history suggests a genetic etiology for the condition in the kindred. At-risk individuals in the family often will be advised to pursue frequent early detection measures. In this context, a genetic test of low or modest clinical validity may not alter the clinical strategy. That is, those who test positive for a known mutation will be advised to continue close surveillance, whereas those with a negative test also may be advised to continue close surveillance, on the assumption that other genes or environmental factors could be the explanation for the family history, or because the condition is common in the general population. Limited clinical utility was a factor in professional recommendations against clinical use of genetic testing for increased risk of melanoma.[52,53] Under the heading of therapeutic choices, the EWG will evaluate research that demonstrates a clinician's response to the information, not data on whether those responses enhance patient welfare. At the same time, the EWG recognizes that clinical response to information may depend on whether there is potential benefit to the patient.

Adherence to the therapeutic or preventive regimen can also be important to improved health outcomes in those circumstances where the regimen is known to be effective. Measurement of adherence or compliance to clinical recommendations is essential to accurately assess both the efficacy and safety of the interventions. Adherence is relevant because many interventions, such as colonoscopy and mammography, must be undertaken consistently and repeatedly over many years to achieve their potential effectiveness. Patient knowledge, understanding, and choice can enhance adherence,[54] but other factors, such as coverage of costs and availability of the recommended measures, may be important as well. It is also appropriate to measure adherence to general population recommendations, e.g., adherence to routine mammography or colorectal cancer screening, for those with negative genetic test results. Theoretically, the net effect of a genetic test in a population could be negative if those at increased risk do not adhere to recommendations for whatever reason, whereas those who are not at increased risk forgo surveillance based on false reassurance. Problems related to adherence and compliance are by no means unique to scenarios involving genetic testing, and understanding both whether individuals are compliant and why they behave as they do is important to the success of any testing program.

Health-related behavior represents an intermediate outcome between testing and impacts on morbidity or mortality. In general, EGAPP does not consider a change in behavior to be sufficient for making a recommendation, except where a behavior is incontrovertibly linked to an outcome, e.g., smoking cessation with lung cancer and overeating with obesity. Nonetheless, a great hope in public health is that behavior change may be an important factor linking testing with changes in health outcomes. Research can assess some health-related behaviors in response to genetic test results relatively soon after testing. Observing adherence over time is important but more difficult. The value of this intermediate outcome depends on the evidence for an association between the health behavior and reductions in morbidity or mortality. Because mammography in women aged 50 years and older reduces the risk of death from breast cancer, increases in mammography rates for older women in response to genetic testing may be advantageous, if linked to appropriate intervals for testing and suitable follow-up and management of tests. For example, if a genetic test provides the impetus for a woman to pursue regular screening, the enhanced adherence should improve health outcomes, whereas increasing frequency of screening in a woman from annually to semiannually is unlikely to yield meaningful health benefits. Demonstrating the actual impacts on morbidity and mortality data will be important because it cannot be assumed that, for example, mammography is equally effective in women at risk by virtue of specific genetic mutations compared with women in the general

population. Furthermore, clinical recommendations for individuals at increased risk may extend beyond what the evidence supports for individuals at average risk. For example, some recommendations for women with strong family histories and *BRCA1/2* mutations encourage mammography at ages younger than 40 years,[55,56] with potential use of magnetic resonance imaging as an adjunct.[57] However, the utility of mammographic screening in women aged 40 years and younger, and those at increased risk of breast cancer, remains controversial.[58] Therefore, adherence to mammography recommendations by younger women is not a good intermediate outcome in the absence of data that mammography in younger women with *BRCA1/2* mutations is associated with decreased morbidity or mortality. That is, it is only speculative that adherence to mammography recommendations in younger women with *BRCA1/2* mutations will lead to reductions in morbidity or mortality. It is reasonable for professional groups to fashion recommendations for health behaviors after genetic testing and for clinicians to recommend such measures to their patients, but ultimately the clinical utility of the test will depend on demonstrating a link between health-related behaviors and reductions in morbidity or mortality for those responding to genetic test information. The value of intermediate outcomes such as adherence to preventive measures is contingent on the strength of the relationship between the intermediate outcome and the primary outcomes of morbidity or mortality. In addition, their value may depend a great deal on the individual patient.

PATIENT OUTCOMES

A premise of the EWG is that it is important to assess the clinical utility of tests before their widespread introduction. The EWG is primarily concerned with changes in morbidity (including quality of life) and/or mortality that occur as a consequence of a genetic test. These outcomes include the benefits and/or harms that accrue to true positives (those with the condition correctly identified), the false positives (those without the condition who are incorrectly identified and potentially managed inappropriately), and false negatives (those with the condition who are not correctly identified and therefore do not benefit from or have the harms associated with management of true positives). True negatives may have improved outcomes as a result of avoiding unnecessary interventions or perhaps worse outcomes due to false reassurance (that is, a belief that a negative genetic test result means they are at no risk from the condition rather than at a general population risk). As in the assessment of nongenetic tests, the central question is whether the use of genetic test information leads to lower morbidity or mortality across the spectrum of patients for whom the test is used, when compared with management for those same patients who are not tested.

Although key to the clinical utility analysis, some of these outcomes may be difficult to assess. Morbidity and mortality may take years or decades to become manifest and may result from clinical events, including additional diagnostic and therapeutic maneuvers, as well as from the impacts of a genetic mutation. Monitoring the health of those tested and not tested can be a long-term enterprise. Genetic factors are usually probabilistic, meaning that adverse consequences are not unique to those with genetic predispositions. For example, most women with breast cancer are not *BRCA1/2* mutation carriers,[59,60] meaning that outcome studies must include a relatively large number of subjects to demonstrate efficacy. To further complicate documentation of morbidity and mortality, it may be difficult to conduct a randomized trial when testing is readily

available in clinical practice. Prevalent presumptions about the short-term and long-term benefits of genetic information may make it difficult to randomize individuals to a no-test group and maintain study subjects' no-test status over time. For some genetic test applications, the impacts on health may be apparent in the shorter term. Pharmacogenetic tests are designed to identify individuals who may be more or less responsive to drugs at standard doses and more or less susceptible to side effects. In some cases, changes in morbidity from genetic testing in the area of pharmacogenetics may be identifiable in a shorter time period than other genetic testing applications. However, even in pharmacogenetics, long-term observation may be necessary for preventive medications and treatment of chronic conditions.

For genetic conditions with frequently fatal outcomes such as cancer, the assessment of mortality rates (e.g., 5 years and 10 years postsusceptibility testing) in those tested and not tested would be important. For other genetic conditions, an evaluation of morbidity through the assessment of benefits and adverse outcomes would be more relevant. Assessing the presence of cancer 5–10 years later in groups with and without genetic testing would provide information on whether preventive or early detection measures are enhanced by testing. For pharmacogenetic applications, an assessment of rates and severity of drug side effects would be important. Intermediate variables such as time to optimal dosing may provide useful information but will generally be inadequate for EGAPP to make a positive recommendation unless this change yields lower morbidity or mortality. Where not obvious, ascertainment of changes in quality of life may also be required.

FAMILIAL AND SOCIETAL IMPACT

As noted, a hallmark of genetic tests for heritable conditions is that they provide some information about the genetic status of relatives within the same kindred.[43] If a woman is found to have a *BRCA1* mutation, one biological parent also is a mutation carrier. In addition, her siblings and children are at 50% risk of being carriers (other than in instances of spontaneous mutations or misattributed paternity). The importance of this finding, and interpretation of risk, is highly dependent on the family history of breast and related cancers and the age of occurrence. Once a mutation carrier has been identified within a family, the cancer risk in the kindred should be determined to assess whether counseling of other at-risk family members about the value of *BRCA* testing is warranted.[61] To the extent that genetic testing could provide benefits, the benefits are multiplied by channeling resources to at-risk relatives. If an individual with colon cancer at a young age is determined to have a genetic variant associated with Lynch syndrome, mutation testing in at-risk relatives followed by preventive measures may be a highly effective strategy to bring genetic testing to those most likely to benefit.[62,63] The implications of similar results for the family members of an older adult may be quite different, however (e.g., for an 80-year-old individual with siblings of similar age, none of whom have colorectal cancer or children). Therefore, an important set of outcomes for some types of genetic testing arise from the impacts on family members that flow from one original test. Relevant questions include: are family members informed of their genetic risk? do family members want to know their risk? how many pursue genetic counseling or clinical evaluations, and how many at-risk relatives pursue genetic testing? Longer term studies might include psychological and behavioral responses to testing in family members and measures of morbidity and mortality. The central point is that outcomes in this context

often must include the effects of testing that cascade through the proband's family.

Societal impacts of the use of a genetic test can be diverse. Cost issues are relevant to any new technology, though cost is not a primary consideration in EGAPP evaluations.[17] A cost-benefit analysis will compare the total societal costs with the total societal benefits in economic terms. Cost-effectiveness analysis is used to compare the costs of competing strategies with the health outcomes obtained.[64] Both of these forms of analysis may be relevant to genetic testing and screening. However, these economic analyses depend on good data on the outcomes noted above and predictable processes of care after testing. Although economic evaluations based on decision models may be informative, the quality of the economic evaluation of a genetic test is only as good as the data and assumptions on which they are based. The literature on economic outcomes of genetic tests should be carefully scrutinized for the quality of the assumptions and evidence used in the analysis.

Genetic testing or screening also may yield outcomes relevant to specific racial or ethnic groups within society. Tay-Sachs disease carrier frequency is high in Ashkenazi Jews, and carrier testing and prenatal testing have been primarily used in this community.[65] Pharmacogenomic testing may also suggest that some racial or ethnic groups are more or less treatable with certain drugs due to the prevalence of different DNA polymorphisms within those populations. Indications for many genetic tests or screening tests include the racial or ethnic background of the patient. Furthermore, research for genes associated with certain conditions will preferentially recruit subjects from racial or ethnic groups with a higher prevalence of the condition. For other conditions, the incidence of certain mutations may vary considerably among population groups. For these reasons, some mutations are associated with racial or ethnic groups in ways that can reinforce group identity and reinforce social stigma. These associations may be seen as beneficial as a stepping stone to prevention and treatment, or they may be seen as reinforcing negative stereotypes. These social attitudes may be important potential outcomes of new genetic tests when their impact on social groups could increase or decrease existing disparities. The measurement of attitudes and discriminatory practices would address these outcomes, although the conduct of this research would be challenging.

At an even broader social level, there are concerns that the contemporary focus on genetics and genetic testing may be steering attention and resources away from other causes of both illness and health disparities.[66–68] It is quite clear that there are environmental and social etiologies for poor health. Given that many health conditions have both genetic and environmental factors, are we focused excessively on the factors that are least amenable to change? Should we be placing more emphasis on the understanding and reduction of environmental triggers of deleterious genes and less emphasis on genetic mechanisms? This general concern about the appropriate use of social resources is rarely relevant to the assessment of individual genetic tests but may be relevant to classes of genetic tests. For example, we should question whether population genetic testing for cardiovascular disease risk will provide helpful information when we already know about large, readily ascertainable, and treatable factors such as smoking, hypertension, and limited exercise. That is, spending individual and social resources on genetic testing for cardiovascular risk in the general population may draw attention and resources from more effective strategies. Although these are important societal resource allocation decisions and the EWG is cognizant of them, these issues are not generally considered as part of the outcomes. Where possible they are considered is as a part of the contextual issues used when translating evidence into recommendations. In the same vein, technology innovators may be drawn to the development of genetic tests rather than tackling innovative ways to reduce smoking or limit exposure to environmental pollutants. Measuring attitudes and behaviors relevant to alternative non-genetic approaches to a disease may be important when strategies emerge using genetic tests in a particular domain. In addition, the expanding area of comparative effectiveness research may be able to shed light on the extent to which testing for particular gene variants adds to assessments based on more conventional factors such as family history or other biomedical test results.

CONCLUSIONS

The EWG has identified a set of outcomes which it will consider for each test to be evaluated and will identify the outcomes that will be examined in the evidence review. Each test under consideration must be evaluated in terms of its suggested indication(s), the clinical context, and the proposed benefits of the knowledge gained through testing. The emphasis of the EWG on morbidity, including quality of life, and mortality is consistent with the approaches of similar groups focusing on other areas of diagnostics and therapeutics. However, "softer" outcomes such as psychological impacts, behavior changes, and social impacts are particularly important in the assessment of many genetic tests and may be weighted accordingly in the EWGs evaluation of specific testing scenarios. Because of the new nature of many genetic tests, data may be limited or nonexistent on many of the outcomes addressed in this article. The lack of reliable data on key outcome measures is likely to lead to a recommendation by the EWG against utilization of the test or a finding that data are insufficient to recommend for or against testing. In these contexts, the EWG will attempt to identify key gaps in the research record to foster studies to fill those gaps. The EWG recognizes the substantial challenges in funding, organizing, and conducting clinical research to address the outcomes of interest in its evaluations. However, the safe and effective use of these sophisticated genetic technologies requires a foundation of evidence to guide clinicians, the health professions, and society.

REFERENCES

1. Guttmacher AE, Collins FS. Welcome to the genomic era. *N Engl J Med* 2003;349:996–998.
2. Burke W, Psaty BM. Personalized medicine in the era of genomics. *JAMA* 2007;298:1682–1684.
3. GeneTests website. Available at: http://www.genetests.org/. Accessed March 20, 2009.
4. About EGAPP. Available at: http://www.egappreviews.org/about.htm, 2008. Accessed November 28, 2008.
5. Schmidt C. Regulators weigh risks of consumer genetic tests. *Nat Biotech* 2008;26:145–146.
6. Genetics and Public Policy Center. Who regulates genetic tests?, 2008. Available at: http://www.dnapolicy.org/policy.issue.php?action=detail&issuebrief_id=10. Accessed March 20, 2009.
7. Genetics and Public Policy Center. FDA Regulation of Genetic Tests, 2008. Available at: http://www.dnapolicy.org/images/issuebriefpdfs/FDA_Regulation_of_Genetic_Test_Issue_Brief.pdf. Accessed March 22, 2009.
8. U.S. System of Oversight of Genetic Testing. A Response to the Charge of the Secretary of Health and Human Services. Report of the Secretary's Advisory Committee on Genetics, Health, and Society, 2008. http://oba.od.nih.gov/oba/SACGHS/reports/reports/SACGHS_oversight_report.pdf. Accessed on January 4, 2009.
9. Evaluation of Genomic Applications in Practice and Prevention. Available at: http://www.egappreviews.org/. Accessed March 22, 2009.
10. CDC Office of Public Health Genomics. About EGAPP. Available at: http://www.cdc.gov/genomics/gtesting/EGAPP/about.htm. Accessed March 22, 2009.
11. What is the EGAPP Working Group? Available at: http://www.egappreviews.org/resources/faq.htm#working_what, 2008. Accessed November 19, 2008.

12. Collaboration Education and Test Translation Program. Available at: http://www.cettprogram.org/. Accessed November 19, 2008.

13. EGAPP Resources. Frequently asked questions, 2008; Available at: http://www.egappreviews.org/resources/faq.htm. Accessed March 22, 2009.

14. U.S. Preventive Services Task Force (USPSTF). Available at: http://www.ahrq.gov/clinic/USpstfix.htm. Accessed March 22, 2009.

15. The Cochrane Collaboration. Available at: http://www.cochrane.org/. Accessed March 22, 2009.

16. Health Technology Assessment International. Available at: http://www.htai.org/. Accessed March 22, 2009.

17. Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. *Genet Med* 2009;11:3–14.

18. Haddow JE, Palomaki GE. ACCE: A model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Burke W, editors. Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease. New York: Oxford University Press, 2004:217–233.

19. CDC Office of Public Health Genomics. ACCE Project, 2007. Available at: http://www.cdc.gov/genomics/gTesting/ACCE.htm. Accessed March 22, 2009.

20. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001;20(suppl 3):21–35.

21. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med* 2007;147:871–875.

22. Tatsioni A, Zarin DA, Aronson N, et al. Challenges in systematic reviews of diagnostic technologies. *Ann Intern Med* 2005;142:1048–1055.

23. Recommendations from the EGAPP Working Group: testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. *Genet Med* 2007;9:819–825.

24. Matchar DB, Thakur ME, Grossman I, et al. Testing for cytochrome P450 polymorphisms in adults with non-psychotic depression treated with selective serotonin reuptake inhibitors (SSRIs). *Evid Rep Technol Assess (Full Rep)* 2007;(146):1–77.

25. Friedman JM, Baross A, Delaney AD, et al. Oligonucleotide microarray analysis of genomic imbalance in children with mental retardation. *Am J Hum Genet* 2006;79:500–513.

26. Watson MS. Current status of newborn screening: decision-making about the conditions to include in screening programs. *Ment Retard Dev Disabil Res Rev* 2006;12:230–235.

27. Cummings S. The genetic testing process: how much counseling is needed? *J Clin Oncol* 2000;18(suppl 21):60S–64S.

28. Claes E, Evers-Kiebooms G, Denayer L, et al. Predictive genetic testing for hereditary breast and ovarian cancer: psychological distress and illness representations 1 year following disclosure. *J Genet Couns* 2005;14:349–363.

29. Hallowell N, Foster C, Ardern-Jones A, Eeles R, Murday V, Watson M. Genetic testing for women previously diagnosed with breast/ovarian cancer: examining the impact of BRCA1 and BRCA2 mutation searching. *Genet Test* 2002;6:79–87.

30. Schlich-Bakker KJ, ten Kroode HF, Ausems MG. A literature review of the psychological impact of genetic testing on breast cancer patients. *Patient Educ Couns* 2006;62:13–20.

31. Aktan-Collan K, Haukkala A, Mecklin JP, Uutela A, Kaariainen H. Psychological consequences of predictive genetic testing for hereditary non-polyposis colorectal cancer (HNPCC): a prospective follow-up study. *Int J Cancer* 2001;93:608–611.

32. Beery TA, Williams JK. Risk reduction and health promotion behaviors following genetic testing for adult-onset disorders. *Genet Test* 2007;11:111–123.

33. Collins V, Meiser B, Gaff C, St John DJ, Halliday J. Screening and preventive behaviors one year after predictive genetic testing for hereditary nonpolyposis colorectal carcinoma. *Cancer* 2005;104:273–281.

34. Decruyenaere M, Evers-Kiebooms G, Cloostermans T, et al. Psychological distress in the 5-year period after predictive testing for Huntington's disease. *Eur J Hum Genet* 2003;11:30–38.

35. Esplen MJ, Madlensky L, Butler K, et al. Motivations and psychosocial impact of genetic testing for HNPCC. *Am J Med Genet* 2001;103:9–15.

36. Gritz ER, Peterson SK, Vernon SW, et al. Psychological impact of genetic testing for hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 2005;23:1902–1910.

37. Meiser B, Collins V, Warren R, et al. Psychological impact of genetic testing for hereditary non-polyposis colorectal cancer. *Clin Genet* 2004;66:502–511.

38. van Oostrom I, Meijers-Heijboer H, Lodder LN, et al. Long-term psychological impact of carrying a BRCA1/2 mutation and prophylactic surgery: a 5-year follow-up study. *J Clin Oncol* 2003;21:3867–3874.

39. Vastag B. Cystic fibrosis gene testing a challenge: experts say widespread use is creating unnecessary risks. *JAMA* 2003;289:2923–2924.

40. Giardiello FM, Brensinger JD, Petersen GM, et al. The use and interpretation of commercial APC gene testing for familial adenomatous polyposis. *N Engl J Med* 1997;336:823–827.

41. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 2008;359:2208–2219.

42. Lyssenko V, Jonsson A, Almgren P, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med* 2008;359:2220–2232.

43. Schulze MB, Weikert C, Pischon T, et al. Use of multiple metabolic and genetic markers to improve the prediction of type 2 diabetes: the EPIC-Potsdam Study. *Diabetes Care* 2009;32:2116–2119.

44. Szabo CI, Worley T, Monteiro AN. Understanding germ-line mutations in BRCA1. *Cancer Biol Ther* 2004;3:515–520.

45. van Dijk S, van Asperen CJ, Jacobi CE, et al. Variants of uncertain clinical significance as a result of *BRCA1/2* testing: impact of an ambiguous breast cancer risk message. *Genet Test* 2004;8:235–239.

46. Green MJ, Botkin JR. "Genetic exceptionalism" in medicine: clarifying the differences between genetic and nongenetic tests. *Ann Intern Med* 2003;138:571–575.

47. Applebaum-Shapiro SE, Peters JA, O'Connell JA, Aston CE, Whitcomb DC. Motivations and concerns of patients with access to genetic testing for hereditary pancreatitis. *Am J Gastroenterol* 2001;96:1610–1617.

48. Fulda KG, Lykens K. Ethical issues in predictive genetic testing: a public health perspective. *J Med Ethics* 2006;32:143–147.

49. Peterson EA, Milliron KJ, Lewis KE, Goold SD, Merajver SD. Health insurance and discrimination concerns and *BRCA1/2* testing in a clinic population. *Cancer Epidemiol Biomarkers Prev* 2002;11:79–87.

50. Shinaman A, Bain LJ, Shoulson I. Preempting genetic discrimination and assaults on privacy: report of a symposium. *Am J Med Genet A* 2003;120A:589–593.

51. Hudson KL, Holohan MK, Collins FS. Keeping pace with the times—the Genetic Information Nondiscrimination Act of 2008. *N Engl J Med* 2008;358:2661–2663.

52. Kefford R, Bishop JN, Tucker M, et al. Genetic testing for melanoma. *Lancet Oncol* 2002;3:653–654.

53. Kefford RF, Newton Bishop JA, Bergman W, Tucker MA. Counseling and DNA testing for individuals perceived to be genetically predisposed to melanoma: a consensus statement of the Melanoma Genetics Consortium. *J Clin Oncol* 1999;17:3245–3251.

54. Fraenkel L, McGraw S. What are the essential elements to enable patient participation in medical decision making? *J Gen Intern Med* 2007;22:614–619.

55. 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.

56. Smith RA, Saslow D, Sawyer KA, et al. American Cancer Society guidelines for breast cancer screening: update 2003. *CA Cancer J Clin* 2003;53:141–169.

57. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin* 2007;57:75–89.

58. Bradbury A, Olopade OI. The case for individualized screening recommendations for breast cancer. *J Clin Oncol* 2006;24:3328–3330.

59. National Cancer Institute. Genetic testing for BRCA1 and BRCA2: it's your choice, 2002. Available at: http://www.cancer.gov/cancertopics/factsheet/risk/brca. Accessed March 22, 2009.

60. Wooster R, Weber BL. Breast and ovarian cancer. *N Engl J Med* 2003;348:2339–2347.

61. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: recommendation statement. *Ann Intern Med* 2005;143:355–361.

62. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. *Genet Med* 2009;11:35–41.

63. Palomaki GE, McClain MR, Melillo S, Hampel HL, Thibodeau SN. EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. *Genet Med* 2009;11:42–65.

64. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.

65. Sutton VR. Tay-Sachs disease screening and counseling families at risk for metabolic disease. *Obstet Gynecol Clin North Am* 2002;29:287–296.

66. Burke W, Khoury MJ, Stewart A, Zimmern RL. The path from genome-based research to population health: development of an international public health genomics network. *Genet Med* 2006;8:451–458.

67. Holtzman NA, Marteau TM. Will genetics revolutionize medicine? *N Engl J Med* 2000;343:141–144.

68. Sankar P, Cho MK, Condit CM, et al. Genetic research and health disparities. *JAMA* 2004;291:2985–2989.