Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

21-Jun-2006

English - Or. English

DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INDUSTRY
COMMITTEE FOR SCIENTIFIC AND TECHNOLOGICAL POLICY

DSTI/STP/BIO(2006)17
Unclassified

Working Party on Biotechnology

Background Paper on the Evaluation of Genetic Tests

Workshop on "Clinical Validity and Clinical Utility"

To be held at the NOWGEN Centre, Grafton Street, Manchester, United Kingdom on 26-27 June 2006

This document is tabled to stimulate debate at the Workshop on " The Evaluation of Clinical Validity and Clinical Utility of Genetic Tests" to be held in Manchester, 26-27 June, 2006. The authors of this document are Dr. M. Kroese and Dr. R.L. Zimmern of the Public Health Genetics Unit, Cambridge, United Kingdom.

For further information, please contact: Ms. Elettra Ronchi; Tel: (+33 1) 45 24 18 28; Fax: (+33 1) 44 30 63 36; E-mail: elettra.ronchi@oecd.org

JT03211048

English - Or. English

INTRODUCTION

1. The Human Genome Project has been a catalyst for an impressive advance in our knowledge of molecular science and the development of novel genomic technologies. This has resulted in the availability of an increasing number of genetic tests. As with any new medical technology there is now international demand that they be appropriately evaluated, and for their results to inform the regulatory framework and decisions about whether or not they should be implemented in healthcare practice.

2. Genetic diseases are conventionally regarded as those inherited according to known and accepted patterns of inheritance and for which the risk to family members is high. They are also often referred to as inherited diseases. The exact definition of a genetic test is still debated and a range of different definitions is in use. The core of the debate is focussed on whether a genetic test is a test for individuals with or at risk of an inherited disorder versus a genetic test being a test based on DNA technology. For the purposes of this paper, a genetic test is defined as one based on the analysis of human DNA using a variety of different technologies.

3. An assay is a method to analyze or quantify a substance in a sample. The term *genetic test* should be regarded as shorthand to describe a particular assay to detect:[1]

a) A particular genetic variant (or set of variants).

b) For a particular disease.

c) In a particular population; and

d) For a particular purpose.

4. Genetic tests may be carried out for a variety of purposes. These include:

- *Diagnostic testing* to confirm or rule out a known or suspected genetic disorder in a symptomatic individual.
- *Predictive testing* to determine the probability of asymptomatic individuals who are suspected of having an inherited disorder developing the clinical manifestations.
- *Susceptibility (or predisposition) testing* to determine the risk or probability that individuals with the genetic mutation will develop a particular disease.
- *Carrier testing* to identify individuals who have a gene mutation for a disorder inherited in an autosomal recessive or X-linked recessive manner.
- *Prenatal testing* to determine during pregnancy whether there is an increased risk of having a child with a genetic condition.
- *Population screening* to identify asymptomatic individuals from within a particular community or a subsection of that community who have an increased chance of having a specific genetic disorder, of carrying a specific genetic predisposition to disease or of being a carrier of a recessive genetic variant.

5. In addition, there is the field of pharmacogenomics, which investigates how genomic variation influences interindividual variability in drug response. Pharmacogenetic tests identify the presence or absence of a particular variant, which can influence an individual's response to a specific drug.

6. There is evidence that diagnostic tests of all types, not just molecular genetic tests, are often implemented without adequate appraisal.[2-4] Diagnostic tests account for a significant proportion of healthcare budgets in the developed world; the potential availability of genetic tests will increase these numbers, and the benefits of further investment will need to be demonstrated. Associated with the increasing number of available tests is the cross border movement of samples for analysis in different countries.

7. Genetic test evaluation methods are still under development but considerable progress has been made.[5-9] The major framework for such evaluations is the ACCE model developed in the US. The framework takes its name from the four components evaluated: Analytical validity, Clinical validity, Clinical utility and the Ethical, legal and social implications of genetic testing.[9] The analytical validity of a genetic test defines its ability to measure accurately and reliably the genotype of interest. Clinical validity defines its ability to detect or predict the presence or absence of the phenotype or clinical disease. Clinical utility refers to the likelihood that the test will lead to an improved outcome and includes financial costs.

8. There is growing international consensus based on practical experience of genetic test evaluation supporting the ACCE framework.[10-12] The ACCE project was completed in 2004. The Office of Genomics and Disease Prevention (OGDP) at the Centres for Disease Control and Prevention, USA has recently launched a new three-year project, the *Evaluation of Genomic Applications in Prevention and Practice* (EGAPP), building on the work of the ACCE project.[13] The great majority of genetic test evaluations undertaken so far have been for rare single gene disorders.

9. Whilst there has been a remarkable increase in our understanding of genomics and the development of new genomic diagnostics, this has not been reflected in the application of these in healthcare. There are several reasons for this and include:

a) The failure to appreciate that health systems now have an evidence based approach to funding and reimbursement decisions for such interventions.

b) Regulators of diagnostic tests and devices do not normally require evidence of clinical validity and utility (unless specific clinical claims are made).

c) Platforms for the assessment of these new diagnostics are rarely available.

d) Lack of understanding of how to establish the clinical validity and utility of these new diagnostics and the standards that may be required of them for effective clinical practice.

10. The information requirements for the marketing, clinical use and regulation of molecular diagnostics can only be provided through a systematic and validated process of test evaluation. Currently this is not performed adequately in any European state. The United Kingdom has initiated a process of genetic test evaluation for molecular genetic tests for rare diseases, which are provided by the National Health Service (NHS). This is the United Kingdom Genetic Testing Network (UK GTN) *Gene Dossier* evaluation framework based on the ACCE programme.[14] In Canada, Germany and Austria, genetic test evaluation has been considered within the context of Healthcare Technology Assessment (HTA).

11. The European Union is sponsoring considerable activity aimed at the harmonisation of regulatory and quality-assurance standards for molecular laboratories. Eurogentest is an example of one such programme and within the Public Health Program of the European Union, the project PHGEN (Public Health Genetics European Network) has recently been funded to focus on the development of policy based on the validity and utility of genetic tests. The Organisation for Economic Cooperation and Development (OECD) has also identified the need for an international framework to standardise genetic testing methods and procedures across borders. This paper will focus on clinical validity and clinical utility because an OECD working group is exploring the subject of analytical validity in greater detail.

12. This report will not address the regulation of genetic tests but it is important to emphasise that evaluation is a technical or methodological issue and should be distinguished from regulation, which is a matter for policy. Regulation and evaluation of tests should be distinguished from both the regulation and quality assurance of laboratories, and the evaluation of new technologies. The regulation of tests can be considered at three levels.[15] These are:

(a) Statutory regulation.

Based on laws and statutory codes.

(b) Regulation by commissioners and payers.

Healthcare payers or commissioners, which will include insurers and governments, will need to obtain the best value from limited funds. They will purchase new tests on the basis of the level of health benefit they provide and their cost-effectiveness.

(c) Professional regulation.

A system of clinical governance, which might include the development of practice guidelines and health provider education.

Disease characteristics

13. There are particular characteristics of the disease that need to be considered when evaluating genetic tests. These include penetrance, genetic heterogeneity and variable expressivity. Penetrance is the probability that someone with a disease-associated genotype will develop the disease. Penetrance includes a time component and is often described in terms of lifetime penetrance *i.e.* the risk of getting a disease during an average lifetime. For example a woman with a mutation in the BRCA 1 or 2 gene has a risk of up to 80% of developing breast cancer.

14. Genetic heterogeneity describes the property that there can be many different genetic causes of the same disease. A particular genetic condition may be caused by more than one gene, *locus heterogeneity*; for example the condition tuberous sclerosis complex can develop due to one of a large number of mutations in two genes, each on a different chromosome. Or by more than one variant within the gene, *allelic heterogeneity*; there are over a thousand different mutations in the CFTR gene that can lead to the development of the disease cystic fibrosis.

15. Variable expressivity describes the situation where people with the same disease-associated genotype experience different types of symptoms and with varying severity. A mutation in the NF1 gene may result in the condition neurofibromatosis 1 but the same mutation may cause one individual to suffer the severe symptoms of the disorder including learning disability whilst their affected relative may only have mild manifestations.

Clinical epidemiology

16. Measures of test performance are well described in textbooks of clinical epidemiology and are based around the concepts of sensitivity, specificity and positive and negative predictive values (NPV, PPV). Analytical sensitivity is the probability that the test detects the specific mutation or those mutations that the test was intended to detect; analytical specificity is the probability that the test does not detect specific mutations or mutations that are not present. Slightly differently, clinical sensitivity is the probability of a positive test result when disease is present. Clinical specificity is the probability of a negative test result when disease is absent. PPV and NPV are calculated for both analytical and clinical validity. In analytical validity, PPV is the proportion of samples with positive test results that have the mutation of interest and NPV is the proportion of samples with negative test results that do not have the mutation of interest. In the case of clinical validity, PPV is the proportion of patients with positive test results that have the disease and NPV is the proportion of patients with negative test results that do not have disease. All these measures may be calculated by presenting the results of the test and disease (or mutation status) in a 2 X 2 table as shown in Table 1. This approach is less suitable for predictive and predisposition testing because it does not capture the effect of time on the "true" disease category.

Table 1 Measures of test performance

		Disease status		
		Yes	No	
Test	Positive	a (True Positives)	b (False Positives)	***a+b*** (Total test positive)
	Negative	c (False Negatives)	d (True Negatives)	***c+d*** (Total test negative)
		a+c (Total with disease)	***b+d*** (Total without disease)	**a+b+c+d**

Sensitivity = a/(a+c) PPV = a/(a+b)
Specificity = d/(b+d) NPV = d/(c+d)

17. The characteristics of a test are critically influenced by the population on which the test is carried out and on the prevalence of disease in that population. The sensitivity and specificity of a test remains constant as the prevalence of disease changes, but the positive and the negative predictive values vary with disease prevalence, especially for tests of low sensitivity and specificity. This property allows the sensitivity and specificity findings from a study carried out on one population to be applied to other populations with different disease prevalence. However this assumption holds only as long as the clinical spectrum of cases in the diseased and non-diseased groups remain the same in the two populations,[16] in other words if there is no spectrum bias or selection bias that might differentially affect the definitions of disease and non-disease.[17; 18] Figure 1 shows the effect of different disease prevalence on test performance. It demonstrates how the calculated parameters change when the test is used in two different populations.

Figure 1. The effect of prevalence on test performance

50% prevalence

		Condition X		
		Present	Absent	
Test	Positive	475	25	***500***
	Negative	25	475	***500***
		500	***500***	1000

Sensitivity = 95%

Specificity = 95%

PPV = 95%

NPV = 95%

2% prevalence

		Condition X		
		Present	Absent	
Test	Positive	19	49	***68***
	Negative	1	931	***932***
		20	***980***	1000

Sensitivity = 95%

Specificity = 95%

PPV = 28%

NPV = 99.9%

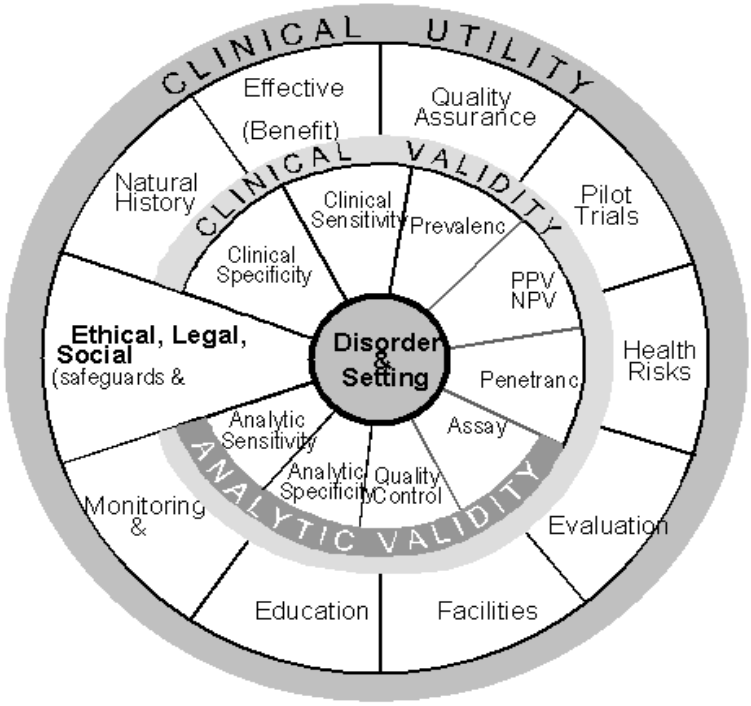
18. In order to establish the analytical and clinical validity of a genetic test, controls need to be included in the assessment. For analytical validity, negative controls are samples that are known not to have the mutation of interest; whilst for clinical validity the controls are individuals that do not have the phenotype of interest according to the disease case definition. It is only through the use of such controls that an accurate assessment of specificity can be made. If possible, the analysis of the results of the tests should be blind to the disease status of participants. This will minimise test review bias.[19] Failure to use controls will critically undermine any performance measures obtained.

ACCE Framework

19. The guiding principle of the ACCE framework is that the evaluation of genetic tests should be an integrated approach including all the domains. This is illustrated in Figure 2. At the core of the framework is the need to define the genetic test as outlined earlier and the healthcare setting in which it is going to be used. Without these definitions the evaluation will produce results of limited value.

20. The ACCE framework is a series of forty-four questions divided into the four domains and following the sequence from laboratory to clinical setting. These are presented in Appendix A. Each question addresses an important part of the evaluation. Generic methodological guidance for addressing each domain is currently not available, although each of the published ACCE reviews and the ACCE website provide examples of different approaches.[20]

Figure 2 ACCE Framework[21]



Analytical validity

21. The analytical validity of a genetic test refers to an assay and defines its ability to measure accurately and reliably the genotype of interest. Therefore this part of the evaluation is concerned with assessing test performance in the laboratory as opposed to the clinic. Explicit specification of the genotype of interest is needed because the estimation of analytic validity is both method- and mutation-specific. This is an important factor when comparing or combining results from different laboratories.

22. The key quantative measures of assay performance for analytical validity are analytical sensitivity and specificity. With DNA based technologies it is possible to achieve analytical sensitivity and specificity close to 100%.[22; 23]

23. Analytical validity can be considered in three parts: *pre-analytical, analytical and post analytical.* Pre-analytical is the stage, which includes obtaining the sample, transport, labelling and entering details on the laboratory database. Analytical is the stage involving preparation of the sample for and carrying out the analysis. Post analytical is the stage including result interpretation and reporting. At each of these stages errors can occur which will affect the analytical validity of a particular test. One US survey reported a sample mix up rate of 0.25% and a survey of European laboratories reported that 20% of participating laboratories made technical or administrative errors in the testing for cystic fibrosis.[24] Follow up reviews indicate that the error rate has declined with the development of external quality assessment programmes.[25]

24. Quality assurance aims to ensure that test results are reliable and reproducible and usually include internal and external control assessments within a quality management framework. For many genetic tests, especially low-volume tests for rare conditions, quality assurance will be the main way of assessing analytic validity. A recent survey on quality assurance of genetic testing services in the European Union

revealed that the participation of laboratories in external quality assessment (EQA) schemes was fragmented and incomplete. Many laboratories did not have any accreditation.[26; 27] These results suggest that the analytical validity of a significant proportion of the genetic tests currently provided by molecular laboratories within the European Union cannot be assured.

Clinical validity

25. Clinical validity defines the ability of a genetic test to detect or predict the presence or absence of the phenotype or clinical disease. It reflects both the clinical sensitivity of the test (the proportion of affected people with a positive test) and the penetrance of the mutations, which the test is designed to identify. Critical to the clinical validity of the test is the understanding of the penetrance of the mutations being tested. The current limits of scientific knowledge means that for some genetic tests all the main causative mutations are not known and this will reduce the clinical validity of the test even if all mutations are tested for. The key causative mutations for a particular disorder can also vary with different populations. For example studies of the clinical sensitivity of the American College of Medical Genetics panel of 25 mutations for cystic fibrosis has estimated that the clinical sensitivity of the panel was 71.9% for non-Hispanic Caucasians, 41.6% for African Americans and only 23.4% for Asian Americans.[28] The clinical sensitivity of this test was limited by the mutations chosen to be included in the panel for testing. The test identified one out of four affected Asian American individuals tested. This raises the question of whether this genetic test has sufficient clinical sensitivity for use in the Asian American population. It highlights the importance of knowledge not only of the penetrance but also the frequency of specific genetic variants in a defined population.[29] This is often not available.

26. It is important to note that a genetic test with near perfect analytical performance may still produce false positives (those who are test positive but do not have the clinical disorder) and false negatives (those who test negative but have the clinical disorder).

27. Three reasons exist for finding a false positive:

a) The test is not one of perfect analytical specificity.

b) The correct genotype has been identified but the phenotype is not present because of reduced penetrance.

c) The disease has been clinically misclassified.

28. The possible consequences of a false positive result include being exposed to unnecessary screening or treatment and social, psychological and economic harm.

29. Whilst there are four reasons for identifying a false negative:

a) The test is not one of perfect analytical sensitivity.

b) Genes other than the one tested for are responsible for the disease (genetic heterogeneity).

c) Variants other than those tested for are responsible for the disease (allelic heterogeneity).

d) The disease has been clinically misclassified.

30. False negative results may delay screening, diagnosis and treatment.

31. The effect of age related penetrance is particularly relevant to the evaluation of predictive genetic tests, because the prospective epidemiological data needed to establish clinical validity for rare disorders is often not available and is unlikely to be obtained. The methodology for evaluating the clinical sensitivity, specificity, negative predictive value and positive predictive value of a predictive genetic test still needs further development. It is likely to involve sophisticated modelling of data. When this information does become available it will add a degree of complexity to the pre and post-test counselling, as the results will be presented as variable probabilities.

32. In contrast to testing for single gene disorders, common complex disorders such as diabetes and hypertension are thought to involve interactions between a number of low penetrance genetic variants and with various environmental factors. The gene variants will therefore be of low predictive value. It is hoped that predisposition genetic tests for such common disorders will have a large public health impact. The current understanding of the gene-gene and gene-environment interactions involved is insufficient to develop a valid predisposition genetic test for these disorders.[30]

Clinical utility

33. Clinical utility refers to the likelihood that the test will lead to an improved outcome. This is the conclusive stage of the evaluation and is dependant but not restricted to the information provided by the other domains. The data obtained from the analytical and clinical validity assessments does not provide all the necessary information that is required to make a decision about the clinical utility of a test.[31; 32] For example what value constitutes an acceptable sensitivity, specificity or positive predictive value will vary from one disorder to another. The interpretation of these results will depend on both the performance characteristics of the test and the value assigned to the benefit of identifying positive cases compared to cases that are missed or wrongly identified. This means that numerical thresholds for the test performance parameters cannot be applied. Each test therefore needs to be assessed for clinical utility on an individual basis.

34. The clinical utility of a test should consider its usefulness as part of an integrated package of care rather than as an isolated investigation. The issues that should be considered include:

- The natural history of the disorder.
- Is appropriate pre and post test counselling available.
- What is the competency of the healthcare practitioner/organisation providing testing.
- Is there an established patient pathway based on test result.
- What effective interventions are available based on test result.
- Does test result alter clinical management.
- Does test result alter prognosis.
- Will the person who has undergone testing have access to post-test follow up care.
- Is the test result information useful for family members.
- What impacts on other aspects of the healthcare system would a testing programme have for example treatment and referral capacity.

- Financial cost of testing programme including pre and post test care.
- Access to testing service.
- Specific ethical, legal and social issues relevant to the test being assessed.
- Political and financial context of assessment.
- In a managed healthcare setting, opportunity cost should be considered.
- How does the new genetic service compare to other healthcare priorities.

35. Consideration of clinical utility requires one to distinguish between whether the genetic test replaces a test of some type already in use, is an additional test to current diagnostics or is an entirely new test. If a genetic test is replacing another test then the assessment of clinical utility must include comparison of cost-effectiveness of the new test and the test already in use. The addition or replacement of a test with a genetic test that is less cost-effective presents an ethical concern that should be explicitly addressed in the evaluation.

36. How benefit is defined and measured can be difficult, as it depends on the views and values of the society /organisation /company /individual on whose behalf this assessment is being carried out. For clinicians, healthcare management organisations and commissioners of healthcare, the focus will be on evidence of measurable health outcomes for example prevention, improved survival, reduction in complications and the cost benefit analysis for these outcomes. Often it is difficult to quantify the outcomes because of the absence of research evidence and predictions have to be made. In which case it is vital that an evaluation programme is put in place to establish prospectively the outcomes of testing.

37. From an individual's perspective, the benefits maybe interpreted very differently. How an individual perceives the value of genetic information will vary and is dependent on a range of sociological and cultural factors. The benefit to an individual of knowing their genetic risk status is difficult to quantify or measure. It is still uncertain whether knowledge of genetic risk will result in improved health behaviour or outcomes.[33-37]

38. The uptake of genetic testing will be affected by individual perception of risk and benefits of testing. This emphases the importance of test counselling. It is also an important consideration for planning the probable impact of a new genetic test in a healthcare service. For example, evidence from the UK showed that 20% of those at risk of Huntington's Disease underwent testing.[38]

39. The final decision on clinical utility is ultimately a subjective one that should be based on evidence. It is expected that the result of an evaluation for a particular test will vary between healthcare systems even if the supporting evidence is identical. There is no "correct" answer for an evaluation of a test. An evaluation must however explicitly consider the key areas outlined otherwise it could be judged invalid.

40. The ACCE model provides some of the elements needed but further work is necessary to develop the detailed framework for clinical utility assessment.

Ethical, legal and social implication of genetic testing (ELSi)

41. This domain is perhaps the most difficult to address as it is wide ranging. Issues include:

- Privacy and confidentiality
- Informed consent
- Data protection
- Insurance
- Eugenics
- Discrimination
- Health inequalities
- Stigmatisation
- Prioritisation decisions in healthcare

42. Assessing ELSi for new specific tests has been very difficult and most of the work so far has instead concentrated on developing general principles or focusing on the use of genetic tests in screening programmes.[20; 39-49] Detailed discussion of the ELSi topics is outside the scope of this paper. ELSi issues specific to a genetic test or how it is going to be used must be addressed in the evaluation and can be considered as part of clinical utility.

Discussion

43. The evaluation framework discussed in this paper is applicable to the evaluation of all types of tests and is not restricted to genetic tests only. Genetic tests may have specific properties or characteristics that will need to be considered and addressed in an evaluation. This also applies to other biomarkers, for example radiological and physiological tests. A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention. The requirement to determine test characteristics in the context of a particular population and for a particular purpose can be generalised to all types of biomarkers and molecular diagnostics. The setting in which the test is carried out is crucial. A test may be highly valid and useful in one clinical context and perform poorly in another.

44. The key barriers to the development of routine evaluation of genetic tests include:

a) Lack of international agreement and support for a standardised approach to the evaluation of genetic tests.

b) Insufficient data for analytical and clinical validity analysis.

c) Limited scope and use of methodologies for the review and meta-analysis of genomic research data for the purposes of genetic test evaluation.

d) Inability to adequately evaluate predictive tests.

e) The absence of an evaluation and policy framework for the assessment of new molecular diagnostics.

45. Quality assurance of genetic testing services is critical to the performance of genetic tests in clinical practice. Research evidence indicates that greater efforts are required internationally to establish the infrastructure necessary.

46. The experience of genetic test evaluations carried out so far indicates the need to develop specific approaches within the overall ACCE framework for tests with different purposes for example predictive and screening tests. The assessment of clinical utility also requires a more detailed and structured framework. This would ensure that the critical areas of this part of the evaluation are explicitly addressed and allow some degree of comparison between tests.

47. The momentum in the development and provision of genetic tests has not been matched to the same degree in efforts to establish the clinical utility of these tests. Due to the lack of this information some tests are being considered in isolation of the disorder, target population and healthcare consequences. This raises serious questions about the appropriateness of such genetic testing. The performance of a particular genetic test and what useful information it can provide will vary depending on which purpose it is used for and which population is being tested. Even if there is valid evidence for the clinical validity of a test, inadequate test information, counselling and no organised healthcare services to provide appropriate follow up and treatment, may result in the misuse and/or misinterpretation of the results by healthcare professionals and the public. This could have serious negative impacts on users and their families.

48. Other consequences of not evaluating genetic tests include the risk of inappropriate investment of new valuable healthcare resources or the diversion of current resources from the provision of clinically effective diagnostics and treatment (opportunity costs) to fund genetic tests of low clinical utility. Evidence of clinical utility will not prevent the use of such genetic tests but will allow purchasers and commissioners to make an informed decision.

49. Commercial groups are marketing genetic tests directly to clinicians and the public. This is not limited to individual countries as the internet provides an effective conduit for advertising and access to testing across the globe. Such development of direct to consumer advertising could further increase the use of genetic tests particularly in the United States.[50] The key public health question is whether such an increase is due to the appropriate use of genetic tests. There is also evidence to show that healthcare professionals feel they need more training in genetics.[50-52]

50. The development of new genetic tests has exceeded our ability to carry out evaluations of them. Tests may therefore be adopted into clinical practice before they have been properly evaluated. Whilst this is not a new phenomenon, it highlights the need for ways of prioritising tests to review and for different levels of evaluation. The UK Genetic Testing Network has found that the information needs for evaluating genetic tests to be used in large populations versus those in very small, well-defined groups are very different. This raises the question of what the balance should be between the degree of detail required to assess test performance, the available resources, and the negative impact of not providing the test in the absence of evidence. In addition, the limitations of our genetic knowledge and technical abilities means that for the moment there are likely to be gaps in the information needed to complete a thorough evaluation of many genetic tests. The inability to identify all disease-related mutations is one example as it makes it difficult to estimate clinical validity.

Conclusion

51. It is important that genetic test evaluation is implemented and an infrastructure is developed to support this activity. Whilst the provision of genetic tests for rare inherited diseases is important, the appropriate implementation of new molecular diagnostics for the common complex diseases such as cancer, coronary heart disease and diabetes will in time have significantly greater impact on population health. The complexities associated with the interpretation of tests for these more common disorders and the use of micro-arrays and proteomics, will be far greater than for genetic tests used in the diagnosis of high penetrance monogenic disorders. Current gaps in methodology and information need to be addressed before such complex evaluations can be undertaken. It is essential that polices and platforms are developed to enable such evaluations to occur.

52. International agreement and support for a standardised approach such as the ACCE framework for the evaluation of genetic tests and biomarkers, is an important first step in developing the capacity necessary to meet future evaluation requirements.

Element	Specific Question
Disorder/Setting	1. What is the specific clinical disorder to be studied? 2. What are the clinical findings defining this disorder? 3. What is the clinical setting in which the test is to be performed? 4. What DNA test(s) are associated with this disorder? 5. Are preliminary screening questions employed? 6. Is it a stand-alone test or is it one of a series of tests? 7. If it is part of a series of screening tests, are all tests performed in all instances (parallel) or are only some tests performed on the basis of other results (series)?
Analytic Validity	8. Is the test qualitative or quantitative? 9. How often is the test positive when a mutation is present? 10. How often is the test negative when a mutation is not present? 11. Is an internal QC program defined and externally monitored? 12. Have repeated measurements been made on specimens? 13. What is the within- and between-laboratory precision? 14. If appropriate, how is confirmatory testing performed to resolve false positive results in a timely manner? 15. What range of patient specimens have been tested? 16. How often does the test fail to give a useable result? 17. How similar are results obtained in multiple laboratories using the same, or different, technology?
Clinical Validity	18. How often is the test positive when the disorder is present? 19. How often is the test negative when a disorder is not present? 20. Are there methods to resolve clinical false positive results in a timely manner? 21. What is the prevalence of the disorder in this setting? 22. Has the test been adequately validated on all populations to which it may be offered? 23. What are the positive and negative predictive values? 24. What are the genotype/phenotype relationships? 25. What are the genetic, environmental or other modifiers?

Clinical Utility	26. What is the natural history of the disorder? 27. What is the impact of a positive (or negative) test on patient care? 28. If applicable, are diagnostic tests available? 29. Is there an effective remedy, acceptable action, or other measurable benefit? 30. Is there general access to that remedy or action? 31. Is the test being offered to a socially vulnerable population? 32. What quality assurance measures are in place? 33. What are the results of pilot trials? 34. What health risks can be identified for follow-up testing and/or intervention? 35. What are the financial costs associated with testing? 36. What are the economic benefits associated with actions resulting from testing? 37. What facilities/personnel are available or easily put in place? 38. What educational materials have been developed and validated and which of these are available? 39. Are there informed consent requirements? 40. What methods exist for long term monitoring? 41. What guidelines have been developed for evaluating program performance?
ELSI	42. What is known about stigmatization, discrimination, privacy/confidentiality and personal/family social issues? 43. Are there legal issues regarding consent, ownership of data and/or samples, patents, licensing, proprietary testing, obligation to disclose, or reporting requirements? 44. What safeguards have been described and are these safeguards in place and effective?

Modified table from http://www.cdc.gov/genomics/gtesting/ACCE.htm December 2005

REFERENCES

1. Kroese M, Zimmern RL, Sanderson S. Genetic tests and their evaluation: can we answer the key questions? *Genet.Med* 2004;6:475-80.

2. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA* 1995;274:645-51.

3. Bogardus ST, Jr., Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA* 1999;281:1919-26.

4. Nierenberg AA,.Feinstein AR. How to evaluate a diagnostic marker test. Lessons from the rise and fall of dexamethasone suppression test. *JAMA* 1988;259:1699-702.

5. Secretary's Advisory Committee on Genetic Testing. Enhancing the Oversight of Genetic Tests, Recommendations of the SACGT. 1-32. 2000. National Institutes of Health.

6. Secretary's Advisory Committee on Genetic Testing. Development of a Classification Methodology for Genetic Test. Conclusions and Recommendations of the Secretary's Advisory Committee on Genetic Testing. 2001.

7. Burke W, Atkins D, Gwinn M, Guttmacher A, Haddow J, Lau J *et al*. Genetic test evaluation: information needs of clinicians, policy makers, and the public. *Am.J Epidemiol.* 2002;156:311-8.

8. Burke W. Genetic testing. *N.Engl.J Med.* 2002;347:1867-75.

9. Haddow J, Palomaki G. ACCE: A Model Process for Evaluating Data on Emerging Genetic Tests. In Khoury M, Little J, Burke W, eds. *Human Genome Epidemiology*, pp 217-33. New York: Oxford University Press, 2004.

10. Sanderson S, Zimmern R, Kroese M, Higgins J, Patch C, Emery J. How can the evaluation of genetic tests be enhanced? Lessons learned from the ACCE framework and evaluating genetic tests in the United Kingdom. *Genet.Med* 2005;7:495-500.

11. Quality & Safety in Genetic Testing: An Emerging Concern; WHO. http://www.who.int/genomics/policy/quality_safety/en/index1.html (Accessed 2.12.2005)

12. Blancquaert, I and Caron, L. Fragile X Syndrome:The role of molecular diagnosis and screening in an integrated approach to services. 1-176. 2002. Quebec, Agence d'Evaluation des Technologies et des Modes d' Intervention en Sante (AETMIS).

13. Office of Genomics and Disease Prevention and CDC. Evaluation of Genomic Applications in Practice and Prevention (EGAPP): Implementation and Evaluation of a Model Approach. http://www.cdc.gov/genomics/gtesting/egapp.htm (Accessed 1-11-2005)

14. UK Genetic Testing Network. http://www.ukgtn.org/dossier.html (Accessed 2.12.2005)

15. Burke W,.Zimmern RL. Ensuring the appropriate use of genetic tests. *Nat.Rev.Genet.* 2004;5:955-9.

16. Brenner H,.Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat.Med.* 1997;16:981-91.

17. Knottnerus JA, Van Weel C. General introduction: evaluation of diagnostic procedures. In Knottnerus JA, ed. *The Evidence Base of Clinical Diagnosis*, pp 1-18. London: BMJ Books, 2002.

18. Mulherin SA,.Miller WC. Spectrum Bias or Spectrum Effect? Subgroup Variation in Diagnostic Test Evaluation. *Ann Intern Med* 2002;137:598-602.

19. Begg CB. Biases in the assessment of diagnostic tests. *Stat.Med.* 1987;6:411-23.

20. Office of Genomics and Disease Prevention and CDC. A Model System for Collecting, Analyzing and Disseminating Information on Genetic Tests. http://www.cdc.gov/genomics/gtesting/ACCE/fbr.htm (Accessed 1.12.2005)

21. Office of Genomics and Disease Prevention and CDC. Evaluation of Genetic Testing http://www.cdc.gov/genomics/activities/fbr.htm. (Accessed 2.12.2005)

22. Palomaki GE, Haddow JE, Bradley LA, Richards CS, Stenzel TT, Grody WW. Estimated analytic validity of HFE C282Y mutation testing in population screening: the potential value of confirmatory testing. *Genet.Med* 2003;5:440-3.

23. Palomaki GE, Bradley LA, Richards CS, Haddow JE. Analytic validity of cystic fibrosis testing: a preliminary estimate. *Genet.Med.* 2003;5:15-20.

24. Dequeker E,.Cassiman JJ. Genetic testing and quality control in diagnostic laboratories. *Nat.Genet.* 2000;25:259-60.

25. Dequeker E, Ramsden S, Grody WW, Stenzel TT, Barton DE. Quality control in molecular genetic testing. *Nat.Rev.Genet.* 2001;2:717-23.

26. Ibarreta d, Bock A, Klein C, and Rodriguez-Cerezo E. Towards quality assurance and harmonisation of genetic testing in the EU. 1-140. 1-9-2003. European Commission.

27. Ibarreta D, Elles R, Cassiman JJ, Rodriguez-Cerezo E, Dequeker E. Towards quality assurance and harmonization of genetic testing services in the European Union. *Nat.Biotechnol.* 2004;22:1230-5.

28. Palomaki GE, FitzSimmons SC, Haddow JE. Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population. *Genet.Med.* 2004;6:405-14.

29. Watson MS, Cutting GR, Desnick RJ, Driscoll DA, Klinger K, Mennuti M *et al*. Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. *Genet.Med* 2004;6:387-91.

30. Haga SB, Khoury MJ, Burke W. Genomic profiling to promote a healthy lifestyle: not ready for prime time. *Nat.Genet.* 2003;34:347-50.

31. Pinsky L, Atkins D, Ramsey SD, Burke W. Developing guidelines for the clinical use of genetic tests: a US perspective. In Khoury M, Little J, Burke W, eds. *Human Genome Epidemiology*, pp 264-82. New York: Oxford University Press, 2004.

32. Evans JP, Skrzynia C, Burke W. The complexities of predictive genetic testing. *BMJ* 2001;322:1052-6.

33. Marteau TM,.Lerman C. Genetic risk and behavioural change. *BMJ* 2001;322:1056-9.

34. Audrain-McGovern J, Hughes C, Patterson F. Effecting behavior change: awareness of family history. *Am.J.Prev.Med* 2003;24:183-9.

35. Bowen DJ, Ludman E, Press N, Vu T, Burke W. Achieving utility with family history: colorectal cancer risk. *Am.J.Prev.Med* 2003;24:177-82.

36. McClure JB. Are biomarkers useful treatment aids for promoting health behavior change? An empirical review. *Am.J.Prev.Med* 2002;22:200-7.

37. Yoon PW, Scheuner MT, Peterson-Oehlke KL, Gwinn M, Faucett A, Khoury MJ. Can family history be used as a tool for public health and preventive medicine? *Genet.Med* 2002;4:304-10.

38. Harper PS, Lim C, Craufurd D. Ten years of presymptomatic testing for Huntington's disease: the experience of the UK Huntington's Disease Prediction Consortium. *J.Med Genet.* 2000;37:567-71.

39. Morrison PJ. Insurance, unfair discrimination, and genetic testing. *Lancet* 2005;366:877-80.

40. Rothstein MA,.Anderlik MR. What is genetic discrimination, and when and how can it be prevented? *Genet.Med* 2001;3:354-8.

41. Parker M,.Lucassen A. Working towards ethical management of genetic testing. *Lancet* 2002;360:1685-8.

42. Godard B, Raeburn S, Pembrey M, Bobrow M, Farndon P, Ayme S. Genetic information and testing in insurance and employment: technical, social and ethical issues. *Eur.J.Hum.Genet.* 2003;11 Suppl 2:S123-42.:S123-S142.

43. European Commission Expert Group on the Ethical, Legal and Social Aspects of Genetic Testing. 25 Recommendations on the Ethical, Legal and Social Implications of Genetic Testing. 2004. Brussels, European Commission.

44. Burke W, Pinsky LE, Press NA. Categorizing genetic tests to identify their ethical, legal, and social implications. *Am.J.Med.Genet.* 2001;106:233-40.

45. Surbone A. Ethical implications of genetic testing for breast cancer susceptibility. *Crit Rev.Oncol.Hematol.* 2001;40:149-57.

46. Godard B, ten Kate L, Evers-Kiebooms G, Ayme S. Population genetic screening programmes: principles, techniques, practices, and policies. *Eur.J.Hum.Genet.* 2003;11 Suppl 2:S49-87.:S49-S87.

47. Melzer D,.Zimmern R. Genetics and medicalisation. *BMJ* 2002;324:863-4.

48. Zimmern RL. Genetic testing: a conceptual exploration. *J.Med Ethics* 1999;25:151-6.

49. Nuffield Council on Bioethics. Genetic Screening: Ethical Issues. 1-115. 1993. London, Nuffield Council on Bioethics.

50. Genetic testing for breast and ovarian cancer susceptibility: evaluating direct-to-consumer marketing--Atlanta, Denver, Raleigh-Durham, and Seattle, 2003. *MMWR Morb.Mortal.Wkly.Rep.* 2004;53:603-6.

51. Burton H. *Addressing Genetics, Delivering Health. A strategy for advancing the dissemination and application of genetics knowledge throughout our health professions*. Cambridge, Public Health Genetics Unit, 2003

52. Emery J, Watson E, Rose P, Andermann A. A systematic review of the literature exploring the role of primary care in genetic services. *Fam.Pract.* 1999;16:426-45.