



The evaluation of clinical validity and clinical utility of genetic tests

Summary of an expert workshop, 26 and 27 June 2006

NGRL MANCHESTER

National Genetics Reference Laboratory

**Dr M Kroese
Dr R Elles
Dr R L Zimmern**

September 2007

This report is the result of work funded by the UK Department of Health.

Dr Mark Kroese is a consultant in public health medicine at the PHG Foundation.

PHG Foundation
Strangeways Research Laboratory
Worts Causeway
Cambridge CB1 8RN
UK
mark.kroese@phgfoundation.org

Dr Rob Elles is Director of Molecular Genetics

National Genetics Reference Laboratory (Manchester)
and Regional Molecular Genetics Service
St Mary's Hospital
Hathersage Road
Manchester M13 0JH
UK
www.ngrl.org.uk/Manchester

Dr Ron Zimmern is Executive Director of the PHG Foundation.

PHG Foundation
Strangeways Research Laboratory
Worts Causeway
Cambridge CB1 8RN
UK
www.phgfoundation.org

The PHG Foundation is an independent charity that works with partners to achieve better health through the responsible and evidence-based application of biomedical science; it has grown out of the Public Health Genetics Unit (PHGU), which operated within the National Health Service from, 1997-2007.

Contents

Executive Summary	3
1 Background	5
2 Workshop Presentations	7
2.1 <i>Introduction and Policy Context</i>	7
2.2 <i>Genetic Testing and the United Kingdom Genetic Testing Network</i>	8
2.3 <i>Genetic Test Evaluation: Current Practice and Development in the United States</i>	11
2.4 <i>EuroGentest</i>	13
2.5 <i>Evaluation of Diagnostic Genetic Tests in France</i>	14
2.6 <i>The National Institute of Genomic Medicine of Mexico</i>	16
2.7 <i>The Evaluation of Clinical Validity and Clinical Utility of Genetic Tests</i>	16
2.8 <i>Current Practice and Developments in Australia</i>	17
2.9 <i>Challenges in Genetic Test Evaluation: Clinical Validity and Clinical Utility</i>	17
2.10 <i>Testing for Rare Diseases in the UK</i>	19
2.11 <i>Genetic Testing in Practice and Prevention: The US Experience</i>	21
2.12 <i>Testing for BRCA: The Canadian experience</i>	23
2.13 <i>Quality in Genetic Testing: A Patient Perspective</i>	26
2.14 <i>Perspective from Industry</i>	26
3 Working Group Session	29
4 Meeting Summary and Considerations for Policy Action	33
5 Acknowledgements	35
6 Appendices	37
Appendix 1 <i>Workshop Expert Advisory Group</i>	37
Appendix 2 <i>Workshop Background Paper</i>	39
Appendix 3 <i>Workshop Programme</i>	59
Appendix 4 <i>Workshop Participants</i>	63

Executive Summary

An international expert meeting, *The Evaluation of Clinical Validity and Clinical Utility of Genetic Tests*, was held at the Nowgen Centre in Manchester on the 26-27th June 2006, under the auspices of the Organisation for Economic Cooperation and Development (OECD). It was organised in the United Kingdom by the National Genetics Reference Laboratory (Manchester) and the Public Health Genetics Unit. The meeting consisted of invited participants from OECD member countries with experience of genetic test evaluation, health technology assessment, healthcare policy and research. Representatives from a total of 17 member countries were present.

The objectives of the meeting were to:

- Share international experience of genetic test evaluation and review current frameworks, mechanisms, and practice in selected OECD member countries.
- Provide a descriptive analysis of the roles and responsibilities of the agencies involved in genetic test evaluation.
- Consider the ACCE framework (analytical validity, clinical validity, clinical utility and ethical, legal and social implications of the test) and whether this methodology could be translated into internationally agreed principles or recommendations.
- Investigate a number of the key methodological and operational issues for the evaluation of clinical validity and clinical utility of human genetic tests.
- Inform policy makers and regulators about the key issues that surround the technical evaluation of genetic tests.

Major methodological and operational issues regarding genetic test evaluations were also identified and explored. These included:

- Lack of data and evidence to assess clinical validity, clinical utility and test outcomes
- A need for the benefits of genetic tests to be better defined
- Further development of the methodology for the evaluation of genetic tests in different clinical contexts and for different indications
- The requirements for post-implementation data collection
- The needs for reviewing and updating genetic test evaluations

The meeting concluded with the development of the following considerations for OECD countries.

Considerations for policy action

1. Developing an agreed framework for genetic test evaluation.
2. Establishing a formal international network approach to genetic test evaluation.
3. Considering what processes and infrastructure are necessary to develop the function of genetic test evaluation in their individual countries.
4. Establishing incentives and accountabilities for performing genetic test evaluation.
5. Developing consensus quality standards for data and evidence to allow thorough clinical validity and clinical utility analysis of genetic tests.
6. Developing consensus guidelines on genetic test evaluation accessible to the public and health professionals.

The opinions expressed in this report are the sole responsibility of the authors and do not necessarily reflect those of the OECD or of the governments of its member countries.

1 Background

- 1.1 In 2000, the Organisation for Economic Cooperation and Development (OECD) organised an international workshop to review developments in the area of genetic testing. The results of this workshop highlighted the need for an international framework to standardise quality assurance procedures across borders. OECD responded to this finding by organising a survey of its member countries in order to assess molecular genetics laboratory quality assurance and proficiency testing practices. The survey was completed in 2004. The survey assessed laboratory practices associated with DNA-based testing, a rapidly evolving sector of genetic testing. The survey showed that in 2004 molecular genetic testing was provided under widely varying conditions and regulatory frameworks across the 18 participating countries. The limited availability of tests, services and expertise within any one country to deal with many rare diseases means that patient samples and data cross national borders. The report concluded with general recommendations for developing guidelines and particularly emphasised that laboratory accreditation to a defined standard is the single most important measure to ensure the quality of genetic tests. Subsequently an expert group drafted guidelines for the quality assurance of molecular genetic tests which after public consultation were adopted by the OECD Council in May 2007.
- 1.2 The 2004 survey report also pointed to the importance of genetic test evaluation, in particular establishing the clinical validity and clinical utility of newly-developed genetic tests.
- 1.3 There are considerable challenges in establishing and interpreting the clinical validity and clinical utility of genetic tests and these have international relevance, particularly for tests which may be subject to cross-border referral. In order to explore and address these challenges, OECD decided to hold an international meeting to be hosted by the United Kingdom in 2006.
- 1.4 The meeting under the auspices of the Organisation for Economic Cooperation and Development (OECD) was organised by the UK National Genetics Reference Laboratory (Manchester) and the Public Health Genetics Unit (Cambridge) and supported by the UK Department of Health which provided funding for both organisations. It was held at the Nowgen Centre in Manchester on the 26-27th June 2006. The objectives of the meeting were to:
 - Share international experience of genetic test evaluation and review current frameworks, mechanisms, and practice in selected OECD member countries.
 - Provide a descriptive analysis of the roles and responsibilities of the agencies involved in genetic test evaluation.
 - Review statements under the ACCE framework (analytical validity, clinical validity, clinical utility and ethical, legal and social implications of the test) and whether these could be translated into internationally agreed principles or recommendations.

- Investigate a number of the key methodological and operational issues for the evaluation of clinical validity and clinical utility of human genetic tests.
- Inform policy makers and regulators about the key issues that surround the technical evaluation of genetic tests.

1.5 A background paper on the evaluation of genetic tests was prepared for the workshop. This is presented in Appendix 2.

1.6 For the purposes of the meeting, the working definition of a genetic test adopted by the OECD expert group on molecular genetics quality assurance was used:

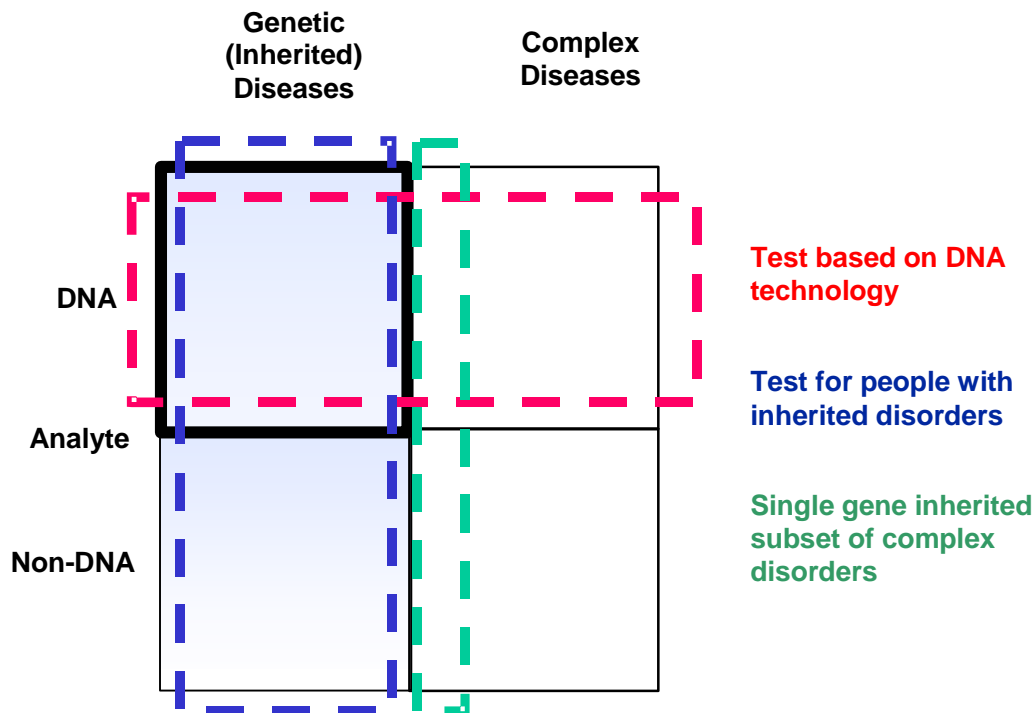
“Genetic testing is testing for variations in germline DNA sequences, or for products/effects arising from changes in heritable sequences, which are predictive of significant health effects.”

2 Workshop Presentations

2.1 Introduction and Policy Context *Dr R L Zimmern (United Kingdom)*

There is considerable debate around the definition of a genetic test. 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. The words *genetic test* are capable of being construed in two different ways: first, as a test for an inherited disorder, and second, as a test based on the analysis of human DNA or other gene-based technologies, such as RNA or chromosomes. In most cases, no attempt has been made to specify whether statements made about genetic tests concern the first or the second meaning. These two meanings are illustrated in Figure 1.

Figure 1 An illustration of the different definitions of “genetic test”



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:

- (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

The difference between an 'assay' and a 'test' should be recognised. An assay quantifies the product; in contrast a test takes into account all the implications when the assay is used for in particular contexts.

The purpose and context of the genetic test is therefore important to define and critical to its evaluation. The major framework for genetic test evaluation 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. Analytical validity of a genetic test defines its ability to measure accurately and reliably the genotype of interest. Clinical validity of a genetic test defines its ability to detect or predict the presence or absence of the phenotype, clinical disease or predisposition to disease. Clinical utility of a genetic test refers to the likelihood that the test will lead to an improved outcome. It was emphasised that the ACCE framework is equally applicable to other forms of molecular diagnostics or biomarkers.

There is an important distinction between evaluation and regulation. Evaluation is a technical exercise whereas regulation is policy driven. Regulation may occur at different levels such as statutory, resource allocation and clinical guideline. Non-statutory regulation is potentially cheaper for society and more responsive to change and public access to accurate information is a powerful regulatory influence.

Tests are becoming more complex, both in terms of the technologies used and in their interpretation. They are also being made more generally available to non-specialists and direct to the public. Physicians have historically been competent in selecting and ordering diagnostic tests but tests involving complex biomarkers may go beyond their current competence. Predictive tests require the acquisition of new skills for their appropriate use.

The complexity of new tests and financial pressures on healthcare funders demand a mechanism to evaluate tests. There is a need to establish technical platforms for test evaluation and to set and agree national and international standards for evaluating tests and methodologies.

Current practice and developments in OECD member states

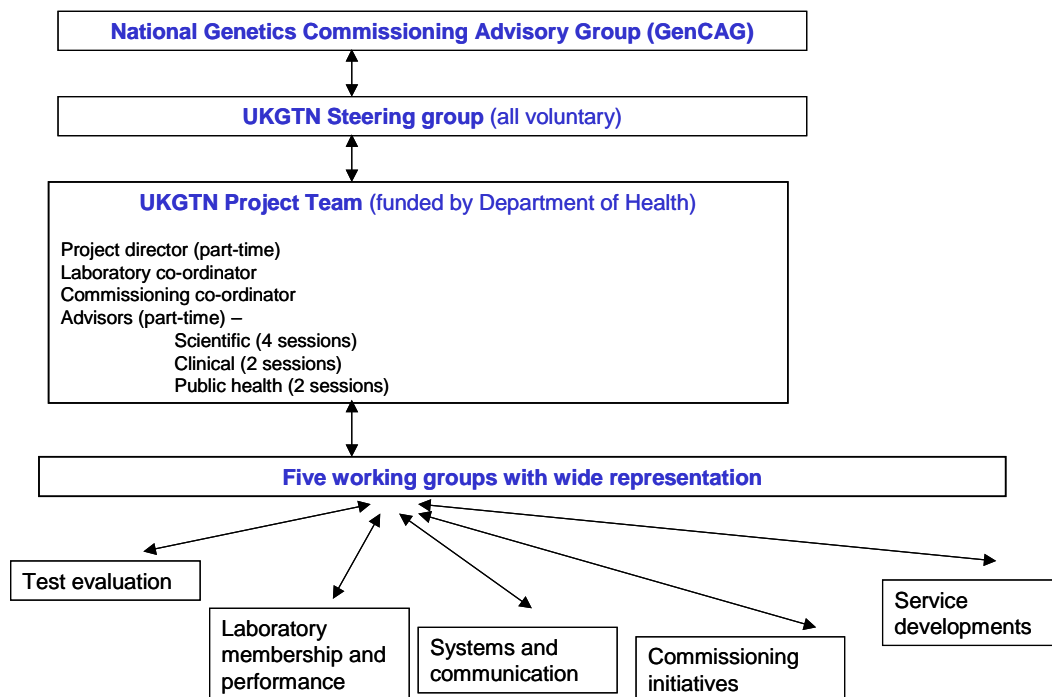
2.2 *Genetic Testing and the United Kingdom Genetic Testing Network* *Professor P Farndon (United Kingdom)*

Clinical molecular genetic testing in the UK began around 1985, initially available through research projects in university and NHS regional genetics laboratories and usually as a result of local interest in particular disorders. During this period of scientific development, and whilst a mechanism for NHS provision was being developed, there were concerns over maintaining continuity of service for particular disorders, ensuring equity of access, a shortage of expert staff and lack of funding for equipment, and lack of resources to develop tests from a research to a clinical base.

Nevertheless, an informal network of service and research laboratories successfully collaborated to develop and provide genetic testing. The members of the network welcomed the establishment of a more formal structure - the UK Genetic Testing Network (UKGTN) - to take forward genetic testing for the UK on a collaborative basis. UKGTN aims to co-ordinate the evaluation, commissioning, funding arrangements and prioritisation of services for inherited genetic disorders. This approach is strongly supported by the UK Department of Health. The UKGTN is a collaborative group of molecular genetic laboratories, their associated clinicians and health service purchasers (commissioners) and is informed by patient groups. Rigorous technical and governance standards are required for laboratory membership.

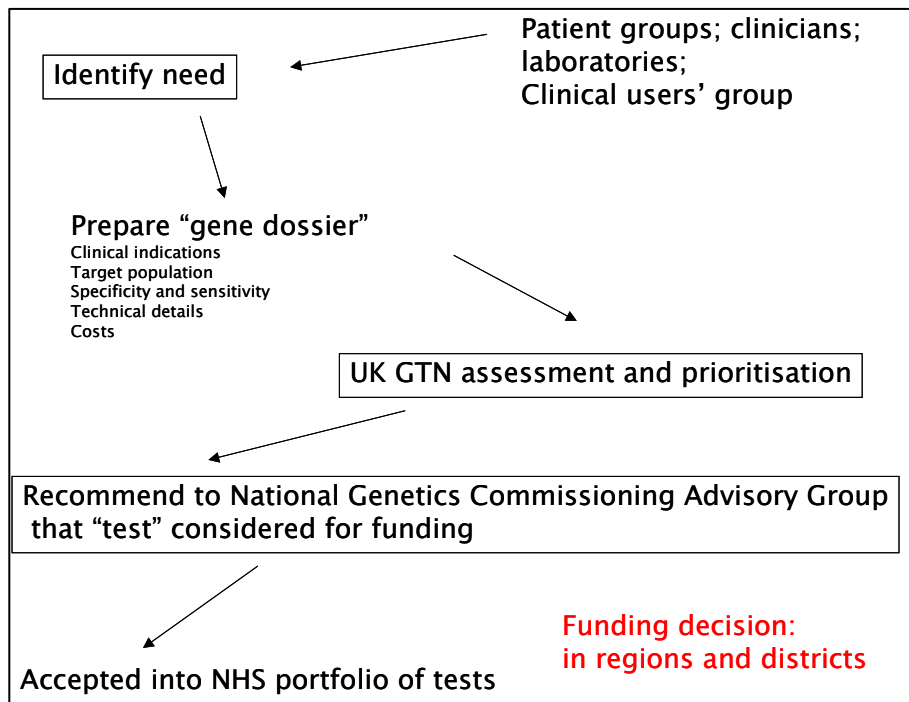
The UKGTN core principle is that of geographical equity. Patients and their families who require genetic advice, diagnosis and management should be able to access a range of expert advice and appropriate tests via local genetics centres which provide the clinical interface and act as a gateway to a coordinated network of laboratory services throughout the UK. The current structure of the UKGTN is shown in Figure 2.

Figure 2 UKGTN 2006



A single mechanism for genetic test evaluation based on the ACCE framework has been developed by the UKGTN. The evaluation process is initiated through the assembly of the 'Gene Dossier' and is linked to the mechanism for adding new genetic tests to NHS service. The infrastructure is presented in Figure 3.

Figure 3 UKGTN infrastructure for the NHS provision of new genetic tests



The UKGTN Directory (April 2005) contains molecular genetic tests for 320 diseases. However genetic tests for some diseases requested through the NHS are required to be sent to non-UKGTN or international laboratories. To date none of the UKGTN directory tests are pharmacogenetic tests. Although the UKGTN system provides a rough guide to assist funding decisions, an independent platform and funding for systematic genetic test evaluation are lacking in the UK.

Future challenges for the UKGTN include:

- Encouraging the appropriate use of genetic testing through educational initiatives and use of testing criteria
- Provision of the development and set up costs for tests
- Adapting to changing NHS funding mechanisms
- The potential tension between a co-operative national network and local decision making and funding priorities

2.3 *Genetic Test Evaluation: Current Practice and Development in the United States* Mr G Palomaki (United States)

In 1997, the Report of the Task Force on Genetic Testing emphasized evidence-based entry of new genetic tests into clinical practice and it defined the criteria for genetic test evaluation to include analytical and clinical validity and clinical utility. This was followed by the 2001 Report of the Secretary's Advisory Committee on Genetic Testing Task Force, which confirmed the importance of ethical, legal and social implications (ELSI) within the evaluation framework. The Centers for Disease Control and Prevention supported a project that developed the evaluation framework further and evolved the ACCE model process. The name of the process reflects the four components of the evaluation which included a consideration of 44 targeted questions. The ACCE system was designed to assess data on DNA-based testing for disorders with a genetic component in a timely manner and provide a clear conclusion. Other features of the process included:

- A "first look" at a broad range of information
- Consideration of data not restricted to published literature
- An *ad hoc* approach to evaluating evidence to extract maximum information from varying sources
- A process to review, extract, analyze, integrate and synthesize data
- Policy recommendations were not intended to form part of the evaluation

The ACCE model project was completed in 2004. Currently the three publicly funded processes for genetic test evaluation in the United States are; the US Preventive Services Task Force (USPSTF), the Evaluation of Genomic Applications in Practice and Prevention Project (EGAPP) and the Collaboration, Education, and Test Translation (CETT) Program.

The USPSTF is an independent panel of experts in primary care and prevention. It is supported by one of the Evidence-Based Practice Centres and develops recommendations on incorporation of interventions into routine primary care. '*Genetic Risk Assessment and BRCA Mutation Testing for Breast and Ovarian Cancer Susceptibility*' was a recent review.

The USPSTF has identified the following challenges in performing genetic test evaluation:

- Limited literature and a rapidly changing evidence base
- The disorders concerned tend to be uncommon
- The desired interventions and outcomes following a test are not clearly defined

EGAPP is a pilot project to establish and evaluate a systematic, evidence-based process for assessing genetic tests in transition from research to practice. It is non-regulatory, independent and non-federal. The project has a 13 member multidisciplinary Working Group and integrates existing processes for the evaluation and appraisal of genetic tests. EGAPP reviews in progress include:

- *Hereditary Nonpolyposis Colorectal Cancer (HNPCC) Testing*
- *Genetic Testing for Detection and Management of Ovarian Cancer*
- *Cytochrome P450 (CYP450) Polymorphisms, Testing in Adults with Depression*
- *UGT1A1 Mutation Analysis in Colorectal Cancer Patients Treated with Irinotecan*

The EGAPP project aims to further develop the ACCE framework for genetic test evaluation and include a consideration of family and societal outcomes. It also aims to shorten the evaluation period. Part of its program will include evaluations of a number of nutrigenomic tests.

The CETT programme is a model program to facilitate translation of genetic tests from research settings to clinical practice. This includes funding for test translation. CETT envisages individual test evaluations through partnerships between a CLIA certified clinical laboratory, a researcher and a patient group with an interest in the test under evaluation. To date thirteen tests have been evaluated or are in progress in a period of just over five months.

There are also privately funded evaluation programmes such as the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC), which carries out evidence based technology assessments. Examples of recent assessments include:

- *Special Report: Genotyping for Cytochrome P450 Polymorphisms to Determine Drug Metabolizing Status*
- *Gene Expression Profiling for Managing Breast Cancer Treatment*

The objectives of genetic test evaluation should be to provide timely, accurate and impartial information on genetic testing. It is also important to recognize the difference between genetic tests in different clinical scenarios and apply appropriate standards for the quantity and quality of evidence on which recommendations can be based.

Full ACCE reviews are time consuming and expensive and should be reserved for situations where testing might be widespread e.g. population screening. However, recent experience shows that for some testing scenarios, a Rapid ACCE review can be completed in less than eight weeks and at a cost of less than \$20,000. This rapid evaluation involves a panel of experts and independent reviewers and is most suitable for topics without a large evidence base.

2.4 *EuroGentest* *Dr U Kristofferson (Sweden)*

EuroGentest is a European Union research funded Network of Excellence aiming to improve the quality and harmonisation of genetic testing across the continent.

Unit 3 of EuroGentest is concerned with public health and clinical genetics. Its work packages comprise development of the concepts of: quality of counseling; clinical validity and clinical utility; evidence based utility; and definitions of genetic testing. The focus is on single gene disorders and Unit 3 does not cover population screening.

The objectives of the clinical validity and clinical utility work package are to establish guidelines for a clinically validated use of genetic testing by:

- Defining clinical utility criteria for genetic testing
- Investigating access to and uptake of genetic testing
- Facilitating international collaboration (ACCE/EGAPP, UKGTN (Gene Dossier), CanGeneTest, OECD and PHGEN)
- Investigating criteria and mechanisms to optimize genetic testing services

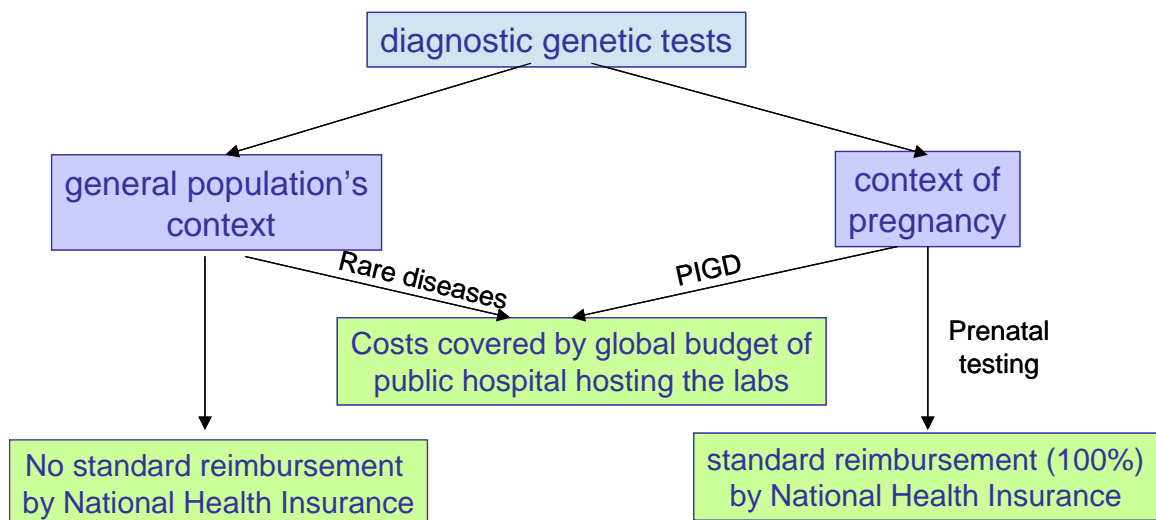
Three thousand five hundred clinical conditions are due to the malfunction of one gene and in 55% of these cases the mutations are identified. However although each of these research findings constitutes a potential clinical test in many cases there are questions over the clinical utility of a molecular genetic test. Six to ten percent of the EU population are affected with a disease where the genetic contribution is the most important factor. This means the provision of genetic testing is already a major challenge for healthcare services. The estimated cost of genetic testing in Germany is 0.07% of the healthcare budget and the aggregate costs of single gene disorders are significant therefore the diagnosis and care of patients at risk of these disorders is a valid public health issue. The work package will consider current definitions of clinical validity and clinical utility and the questions that need to be addressed in these domains of the genetic test evaluation. It is important to consider the purpose of a test during an evaluation which may include diagnostic, predictive or reproductive indications. It is acknowledged that for a majority of single gene disorders the key data does not exist; for example an accurate estimate of the prevalence of the condition. It is difficult to establish the financial and emotional cost and benefits of testing and any evaluation of a genetic test for a rare single gene disorder will be dependent on the clinical context.

2.5 Evaluation of Diagnostic Genetic Tests in France Dr C Carbonneil (France)

Since 2004, three agencies have been involved in the evaluation of genetic tests in France. The *Agence de la Biomédecine* is concerned with legal and ethical aspects regarding human genetics and accreditation of genetic test performing centres and practitioners. The second is the AFSSAPS (the French Health Products Safety Agency) which oversees inter-laboratory quality assessment of tests including medical genetic tests and commercialized diagnostic tests. The third agency is the *Haute Autorité de Santé* (HAS) which develops clinical practice guidelines and performs health technology assessments which include genetic tests.

The provision of genetic tests and their cost reimbursement process is outlined in Figure 4.

Figure 4 Reimbursements of genetic tests in France in 2006



A single framework exists for the evaluation of all diagnostic tests (including genetic tests) and is composed of the assessment of analytical validity, clinical validity, clinical utility and consideration of the target population. The clinical utility assessment includes impact on diagnostic or therapeutic management and prognosis, benefits to patients and relatives and the impact on the health economy.

The assessment is based on data in the published literature and discussions by a multidisciplinary working group briefed to consider the evaluation in the context of the French health system. Reports require approval by the HAS committee for assessment of medical and surgical procedures.

HAS has carried out evaluations of the following genetic tests:

- Factor V (Leiden mutation)
- Prothrombin (FII 20210G>A mutation)
- Haemochromatosis (HFE1 C282Y mutation)
- Haemochromatosis (HFE1 other mutations)
- HER2Neu
- Microdeletions of the Y chromosome
- Down syndrome screening

Whilst the process for genetic test evaluation in France is still evolving, the experience so far has revealed the following difficulties and barriers:

- Evaluation is often requested for non-validated mutation-based tests
- Lack of data regarding validated mutations
- Lack of published data on clinical validity
- Lack of published data on clinical utility
- Limited possibility of using evaluations performed in other countries

Many non-validated mutation-based tests exist, for example the clinical validity of the C282Y mutation in association with haemochromatosis has not been established.

There are certain risks with regard to the timing of an evaluation. If an evaluation is performed too early in the development of a genetic test, there will be a lack of supporting data and other information resulting in the premature rejection of a valid and efficient test. Alternatively, if the evaluation is performed too late, there might be sufficient data and other information but it will be very difficult to stop the use of an invalid and inefficient test, which is already being implemented. In both situations there is a lack of benefit to the patient population and results in a waste of resources.

The European network for Health Technology Assessment, EUnetHTA, coordinates the efforts of 27 European countries including 24 Member States of the European Union in evaluating health technology in Europe.

During the first 3 years of existence (2006-2008) EUnetHTA aims at developing an organisational framework for a sustainable European network for HTA along with practical tools to fill into this framework to ensure timely and effective production, dissemination and transfer of HTA results into useful policy advice to the Member States and EU. Initially, the EUnetHTA project is being co-financed by the European Commission (DG Sanco) and contributions from network members.

Tools for sharing information about emerging technologies (including all types of diagnostic test) are being developed. A risk is that without effective HTA both research and industry will lose credibility and funding. A relationship needs to be established between HTA, professional bodies and scientific groups so that these bodies can request an evaluation and serve on working groups.

2.6 *The National Institute of Genomic Medicine of Mexico*
Dr G Jimenez-Sanchez (Mexico)

In Mexico genetic testing is focused around the neonatal screening programme with DNA based testing being offered mostly by research laboratories in the context of research programmes associated with specific conditions.

The National Institute of Genomic Medicine of Mexico was established as a new initiative in 2004, as the national reference centre for genomic medicine with a \$90m programme. It is a partnership between governmental health and university authorities. Its goal is to conduct translational research and training programmes in genomic medicine. Current areas of research are clustered around common public health problems in Mexico and include obesity, cardiovascular disease, cancer, pharmacogenomics and population genomics; recognising that Mexico includes 65 ethnic groups. Accordingly haplotype mapping in these population groups has been launched. The second emphasis of the centre is on genomics as an economic development activity. Early results show significant genotypic differences between the Mexican ethnic groups and Asian and Caucasian index populations. These differences translate to significantly different risks, for example of an adverse drug reaction following azothioprine prescription, which is stimulating national clinical policy discussions.

The Institute consists of research and teaching, ELSI and four high capacity core technology centres. These are sequencing and genotyping, expression analysis, medical proteomics and supercomputing. It has an important role in developing international collaboration with other genomic centres particularly in Latin America but including biobanking groups such as the Wellcome Trust and the World Health Organisation.

2.7 *The Evaluation of Clinical Validity and Clinical Utility of Genetic Tests*
Dr G Novelli (Italy)

National professional guidelines for genetic testing were established in 1998 and were followed by governmental recommendations for genetic services in 2004 and expert group guidelines for evaluating the analytical and clinical validity of tests in 2005-6. All these efforts focussed on single gene disorders (www.pnlg.it). The approach in Italy has been through the development of evidence based clinical guidelines for example cystic fibrosis and myotonic dystrophy. This process includes a large multidisciplinary group of both national and international experts.

The methodology for guideline development has included the formation of international expert groups including clinicians and patients. Guidelines are drafted after a literature search and submitted for approval to relevant scientific associations and the government.

The genetic investigations associated with infertility, specifically variants in the KAL1 gene were used as a model. The outcome was agreed and referral criteria and a care pathway were developed. The role of CFTR analysis in infertility was similarly assessed and guidelines produced which will be linked to health procedure reimbursement decisions. As a final example guidelines for genetic testing in myotonic dystrophy were approved by the Italian Health Ministry and a standard care pathway has been agreed.

2.8 *Current Practice and Developments in Australia*
Professor R Trent (Australia)

Australia has a federal and tax based national health provision. Most genetic services are the individual responsibility of the six States and two Territories. The Commonwealth (e.g. central) Government funds most cytogenetic tests and six DNA based genetic tests. Test evaluation and standards which are Commonwealth responsibilities do not generally influence the provision of genetic services. The following agencies have a role in genetic test evaluation:

- Medical Services Advisory Committee (MSAC) - assesses and makes recommendations for funding of all diagnostic technologies. At this stage it has no specific process for DNA based genetic tests.
- Therapeutic Goods Administration (TGA) has a role in regulation and in 2007 new regulations covering *in vitro* diagnostic devices will come into force. The TGA considers DNA tests as in the moderate/low risk category for the public but recognises that a genetic test outcome may carry a high personal risk.
- National Pathology Accreditation Advisory Council (NPAAC) - develops standards and guidelines for laboratory testing.
- National Association of Testing Authorities (NATA) - accredits laboratories after an assessment of competence.

The Human Genetics Advisory Committee, part of the National Medical Research Council, is revisiting the impact and issues associated with genetic testing. Efforts are underway to improve the coordination of the evaluation and provision of genetic tests nationally. The difficulty of categorising tests as predictive or diagnostic from the point of view of assigning risks and devising an appropriate level of regulation is recognised.

Evaluation of clinical validity and clinical utility of genetic tests

2.9 *Challenges in Genetic Test Evaluation: Clinical Validity and Clinical Utility*
Professor W Burke (USA)

The components of the ACCE genetic test evaluation are:

- Analytical validity - accuracy and reliability of measurement of genotype of interest.
- Clinical validity - accuracy of detection or prediction of phenotype, clinical disease or predisposition to disease
- Clinical utility - likelihood of improved outcome from test use
- 'ELSI' - ethical, legal and social implications of the test

The concepts of analytical and clinical validity and clinical utility are interconnected and disaggregating these parameters for the purpose of a genetic test evaluation may be difficult. Achieving clinical utility for a genetic test requires actions based on the results of the test, which in turn provides a health benefit. For example, when the test for thiopurine methyltransferase (TPMT) activity in predicting toxicity in chemotherapy for children with leukaemia was considered by the FDA, the clinical utility questions that needed to be addressed included:

- What dose adjustments should be made based on genetic test results?
- How many adverse events will be prevented?
- Will under-treatment of heterozygotes occur?

The distinction between an assay and test proposed by Ron Zimmern is valuable. An assay relates to a genetic variant whereas a test is designed to establish or exclude the presence of a disease in a defined population and for a specific clinical purpose. Extrapolating from one population to another is difficult. For example, the discovery of the BRCA1 and 2 genes relied on (unusual) families with multiple affected persons. The lifetime risk of cancers in BRCA1 or 2 mutation carriers is between 11-65%. However the risk of breast cancer in both BRCA mutation carriers or non-carriers is also independently dependent on environmental factors. These complicating factors make it difficult to acquire high quality evidence on precise risk and therefore on the impact of testing and interventions.

As might have been predicted, given this complexity, the estimates for the risk of breast and ovarian cancer if a woman is BRCA 1 or 2 mutation positive varied depending on the nature of the population from which evidence was obtained, for example multiple case families versus population based. The informativeness of testing for first-degree relatives also varies depending on whether a mutation is identified in the proband. The clinical utility of testing for BRCA 1 and 2 is the reduction of cancer mortality and morbidity through prophylactic surgery, screening and prophylactic drug treatment.

It is a challenge to obtain high quality evidence in medical genetics because of the difficulty in achieving adequate numbers for controlled trials and the unsuitability of randomized trials for many questions. The reasons for this include the nature of the intervention e.g. some interventions, like mastectomy, cannot feasibly or ethically be assessed in a randomized controlled trial. In addition, if presumptive benefit is high, withholding treatment is perceived as unethical. In the case of BRCA 1 and 2 testing the clinical utility questions about intervention for women who have these mutations are: the acceptability of prophylactic surgery, the effect of early mammography and breast MRI on breast cancer mortality, the long-term efficacy of tamoxifen and the effects of different BRCA 1 and 2 mutations for each of these indicators.

Much treatment (in genetics, but also in clinical medicine more generally) is based on presumptive benefit and not based on randomised controlled trials.

The difficulties in assessing clinical utility are even greater when genetic risk is moderate, e.g. in the case of genetic susceptibility for common complex disorders. For example, in the case of thrombophilia the objectives of testing could include:

- The provision of informed choices related to OC / HRT use
- Management of thrombosis risk in pregnancy / surgery
- The identification of rare high-risk individual (such as FVL/PT or homozygote for FVL)

However a recent UK Health Technology Assessment “*Screening for thrombophilia in high-risk situations*” (TREATS Study) concluded that:

- There were limited data for the construction of a cost-effectiveness model
- Selective screening (based on prior venous thrombosis/family history) was more cost-effective than universal screening
- Large prospective studies were needed to refine risks and provide evidence of screening benefits
- The relative value of thrombophilia screening is still to be established

It is also important to consider the benefits of genetic information in the absence of improved mortality and morbidity. These include the improvement in the healthcare process e.g. the end of a diagnostic odyssey and allowing well-informed supportive care and benefits to the well being of the family and carers as well as the patient themselves.

2.10 *Testing for Rare Diseases in the UK* *Dr F Stewart (UK)*

The reasons for performing a genetic test include diagnosis, treatment, prognosis, presymptomatic screening and genetic risk assessment. The clinical utility questions that need to be addressed for each include:

- **Diagnosis**
Can a definitive diagnosis be made for certain by any other method including clinical examination by an expert? Will a molecular diagnosis remove the need for other expensive or invasive tests?
- **Treatment**
Will a specific molecular diagnosis affect treatment?
- **Prognosis and Management**
Will a specific molecular diagnosis affect treatment? Is there evidence in the disease in question that a specific molecular sub-type will affect prognosis and management to a significant extent? In other words – will the result significantly affect the lifestyle choices of the patient or the family?

- **Presymptomatic Screening**

Will a positive molecular result accurately predict future disease and alter management? Will a negative molecular result be definitive (i.e. further tests will not need to be carried out)?

- **Genetic Risk Assessment**

Will molecular diagnosis in the affected person reduce the need for tests in the rest of the family? Will molecular diagnosis resolve the mode of inheritance (e.g. HMSN)? Will molecular diagnosis provide a means of pre-natal diagnosis or carrier detection? Will molecular diagnosis allow pre-symptomatic testing for other family members?

The differences in the clinical utility of genetic testing are highlighted by comparing genetic testing for neurofibromatosis types 1 and 2. In the case of neurofibromatosis 1, the diagnosis is made clinically and genetic test results do not affect treatment or clinical outcome. As the clinical signs for this disorder are evident in early life there is no role for predictive testing. Testing may in rare circumstances have a role in genetic risk assessment for example prenatal diagnosis or in atypical cases. For neurofibromatosis 2, the situation is different as affected individuals may have little in the way of skin lesions and the condition is associated with intracranial and other tumours as well as deafness. A genetic test is important to make the diagnosis in 'at risk' individuals and influences treatment and management. In the UK there is a clinical protocol for the follow-up of individuals with neurofibromatosis 2.

In the UK, a process called the "gene dossier" has been developed by the UK Genetic Testing Network for the evaluation of genetic tests to be provided by the National Health Service. The gene dossier evaluation is based on the ACCE framework. The UKGTN has established a subgroup specifically to assess gene dossier submissions. The membership of this group includes clinicians, clinical scientists, commissioners, public health specialist and representatives from patient groups and the Department of Health. The group may request further information or clarification from applicants. Once a gene dossier has been approved by the working group, it is considered by the UKGTN Steering group for approval and presented to healthcare commissioners (purchasers) for consideration. The assessment of gene dossiers is a transparent process for evaluating genetic tests and all stakeholders are represented in the assessment process. A rejection of a gene dossier may be on the basis of lack of clear utility and may lead to the dossier being referred back to the originator for further work or clarification. Difficulties in the process include the limited evidence base for new genetic tests, and technical limitations. Dossiers may also be rejected if a cheap and effective alternative test is available or if the test result lacks influence on the risk assessment or management of the condition. The gene dossier may be revisited after a period and may be subject to a workshop approach. About 380 tests are on the UKGTN web directory of tests.

The UKGTN experience so far has revealed the following points:

- The disorder and the clinical setting for testing needs to be fully understood and is critical to the evaluation process
- The clinical utility is ultimately a subjective decision but it is the most important domain in the final decision-making process
- The clinical context for testing needs to be clearly defined and include the criteria indicating testing.

The major challenges to genetic test evaluation identified include:

- The evidence base is limited for new genetic tests
- A lack of knowledge and technical limitations means only a proportion of mutations can be identified for a condition by any test
- The prevalence of the disease in the target population may not be known
- Financial data are difficult to obtain

2.11 *Genetic Testing in Practice and Prevention: The US Experience* Professor A Berg (United States)

The US Preventive Service Task Force (USPSFT) is an independent panel of experts in primary care and epidemiology, which conducts rigorous assessments of the scientific evidence on the effectiveness of clinical preventive services. The Task Force then develops recommendations for clinicians in primary care based on their assessment of the evidence for benefit. Any gaps in evidence are also highlighted to inform future research. The scope of its assessments includes primary, secondary and tertiary preventive services. In the area of genetic test evaluation it has or is in the process of performing the following evaluations:

- *Genetic Risk Assessment and BRCA Mutation Testing for Breast and Ovarian Cancer Susceptibility* – September, 2005
- *Screening for Hereditary Haemochromatosis* – review in progress

The USPSFT approach is most appropriate if the condition is common and serious, the intervention and its outcomes are clearly defined, there is a substantial body of evidence including cost effectiveness data and the evidence base changes slowly. The specific challenges of using evidence-based methods for genetic test evaluation include:

- Many conditions are uncommon or rare
- Interventions and clinical outcomes are not well defined

- Tests have inadequate sensitivity and specificity in unselected populations, with poor predictive value
- Tests are proposed and marketed based on descriptive evidence and pathophysiological reasoning, with no clinical trials
- An overlay of advocacy from industry and patient interest groups

The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) project is sponsored by the US Centers for Disease Control. It is an independent non-federal project. It is multidisciplinary and its goal is to establish and evaluate a systematic and sustainable mechanism for pre- and post-market assessment of genomic applications in the United States. It aims to achieve this through an evidence-based approach, which is transparent and publicly accountable. None of EGAPP's methods or processes are final but should be considered as a description of work in progress.

The EGAPP process starts with topic selection, which involves developing brief summaries for possible topics including details on the health burden and any clinical practice issues. These are reviewed and scored, with the results being considered by the project steering group in making its final decision. The evaluation is then performed by AHRQ Evidence Based Practice Centers and includes review of the quality of individual studies, strength of evidence, analytical validity and clinical validity. EGAPP members serve on the Technical Expert Panels for these reviews. For the purposes of evaluating genetic tests the EGAPP group have identified the following categories of outcome:

- Diagnostic thinking and health information impact
- Therapeutic choice
- Patient outcomes impact
- Familial and societal impact

In addition to the major reviews, the EGAPP group has developed a process of fast track reviews for tests, which have a limited evidence base. These reviews are limited in scope and depth and are comparable to technology assessments.

The EGAPP project is planned to last 3 years (currently in its second year) and its products will include 3-5 major reviews, 2-3 fast-track reviews and the development of methods for the evaluation of tests.

An example of an EGAPP evaluation is testing for CYP450 in adults with non-psychotic depression before treatment with selective serotonin reuptake inhibitors (SSRIs). The key questions for this evaluation are:

- Does testing improve outcomes?
- What are the test characteristics?
- Can the test result be correlated with metabolic indicators, efficacy and adverse effects?

- What are the effects of the test on management, clinical outcomes and decision-making?
- Are there harms associated with testing?

Preliminary findings of this evaluation include that there is some data on sensitivity and specificity, no studies linking testing to clinical outcomes, the studies that have been completed are small and poor quality cohort studies, there have been no comparisons of alternative testing strategies and there has been a failure to account for all relevant genotypes.

The experience from the EGAPP process so far has revealed that there are significant gaps in the understanding of the prevalence of the genotype of interest in the general population, penetrance of the genotype, lack of clinical trials comparing testing and intervention, lack of assessment of all relevant outcomes of genetic testing and there is limited data on cost and feasibility for specific genetic tests.

In conclusion, there are a large and growing number of tests marketed in the United States to consumers and clinicians. There is also a national attitude that more is always better and that technology is always good. This occurs in an environment that is hostile toward regulation. Genetic tests have the potential for benefits and harms but evidence is limited.

2.12 *Testing for BRCA: The Canadian experience* *Dr I Blancquaert (Canada)*

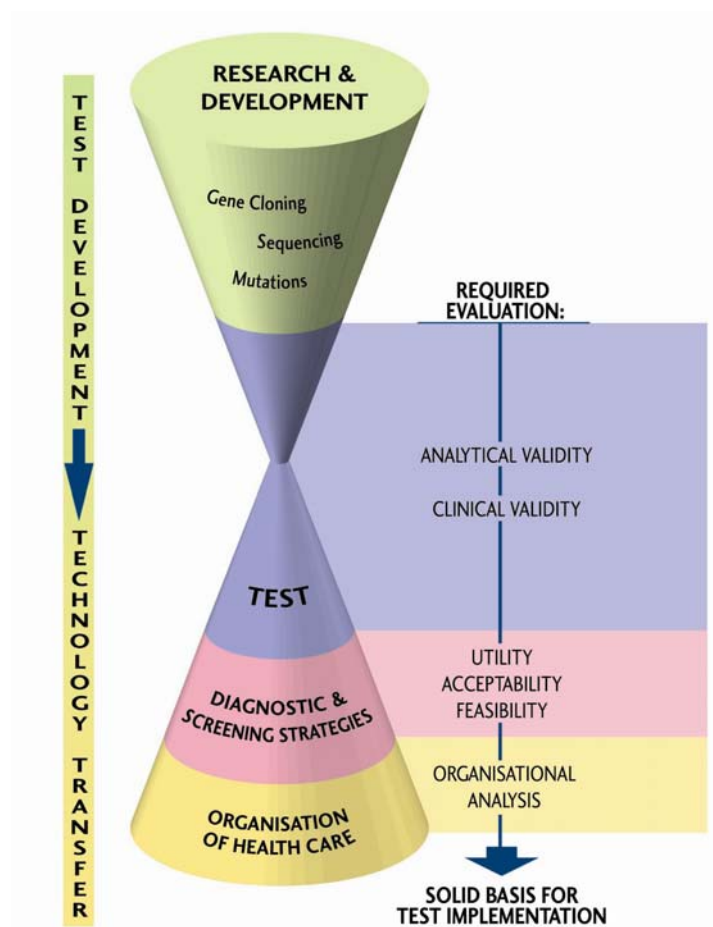
In Canada, the recent trend has been towards supporting evidence-informed policy, with transparency and accountability of the process. The main methodology for this has been health technology assessment (HTA). Since 2004, there has been a concentrated effort to develop HTA capacity, co-ordinate assessments and integrate HTA into an efficient technology transfer process. Canada has one national and three provincial HTA organizations. HTA is defined by INAHTA (International Network of Agencies for HTA) as:

“a multidisciplinary field of policy analysis. It studies the medical, social, ethical and economic implications of development, diffusion and use of health technologies.”

HTA is an evolving methodology and is not limited to systematic reviews and meta-analysis but includes evidence synthesis and contextual analysis. HTA can be considered as comprising both assessment and appraisal. Once the question has been framed, assessment focuses on systematically collecting the evidence, and on performing a critical analysis and synthesis of the evidence. Appraisal requires a contextual analysis and leads to the formulation of recommendations. HTA can provide a comprehensive and balanced perspective of the topic assessed. It is a key to informed decision making.

The HTA framework for genetic tests is shown in Figure 5. Each dimension involves assessing a number of different parameters. Some are similar to those in the ACCE model; others are not labelled in the same way. Ethical, legal and social issues are not mentioned as a separate parameter. In this framework these issues are expected to emerge at all stages of the analysis and need to be integrated accordingly. Also, economic analyses can be performed for each dimension or globally. This comprehensive and multidisciplinary approach to evaluation overlaps to some extent with health services research, policy analysis and public health when organisational issues are considered.

Figure 5 HTA framework for genetic tests



Blancquaert et al., 2001

Important characteristics of the framework include: broad parameters which are not limited to a defined set of questions, use of quantitative and qualitative evidence, evidence combined across multiple dimensions and an emphasis on the critical analysis of all evidence.

Current issues with genetic test evaluation can be illustrated using BRCA testing as an example. Genetic heterogeneity and genotype/phenotype correlation are key characteristics for any genetic test, as they will directly influence clinical validity. There are no studies that have been developed specifically to assess the clinical validity of BRCA testing and, as a result, epidemiologic quality requirements for studies on diagnostic test performance are rarely met. For the hereditary breast and ovarian cancer syndrome, there is no consensus on a phenotypic definition which could serve as reference or gold-standard to assess clinical validity. Failure in study designs to include controls means clinical specificity is usually not measured. A range of different molecular techniques are used for testing and these have evolved over time. The lack of standardised protocols and the variability in selection criteria and testing indications hamper comparisons across studies and prevent pooling of results into summary measures of clinical validity.

In order to address these major obstacles to genetic test evaluation, it is necessary to:

- Define parameters
- Define what constitutes good quality evidence for each parameter
- Agree on criteria for sufficient evidence for a given use of a test
- Agree on phenotype definitions
- Develop concerted research efforts including the development of quality standards and international/regional collaboration
- Develop mechanisms for post-implementation data collection and timely re-evaluation of the knowledge base

In conclusion, collaboration amongst various disciplines is essential to develop and implement genetic test evaluation; particularly between genetic and evidence base medicine specialists, to set quality standards and improve knowledge transfer. Within the wider genetics community there is also a need to develop data sharing, consensus building, collaborative research projects and post-implementation coordination.

2.13 *Quality in Genetic Testing: A Patient Perspective*
Mr A Kent (United Kingdom)

When it comes to genetic tests, a patient wants to know “does it do what it says on the tin?” A test is part of a process and the “quality” of testing relates to the experience of the whole process. For the patient genetic testing is important and is not merely a curiosity driven event. When a patient is considering undertaking a genetic test, they may be anxious, stressed, uncertain and will be hoping that the outcome will be favourable to their expectations.

From a patient’s perspective the dimensions of quality for a genetic test include that: it is available in a timely, user-friendly, patient and family focussed manner and that the information it provides is robust and technically reliable. In addition, the provision of medical information should be honest and it is part of a process with established outcomes. It is proposed that there is alternative genetic test ACCE framework from the patient perspective.

Aavailable
Comprehensible
Compassionate
Equitable

2.14 *Perspective from Industry* *Dr F Lasky (United States)*

There is no clear definition of clinical utility, although the term is used liberally in various fields, such as genetics, pharmacogenetics and biomarkers. Genetic information can add value to patients’ decisions related to both health and to life choices. The concept of “clinical utility,” when applied only to the diagnostic or therapeutic aspects of a test, does not adequately satisfy patients’ concerns when faced with diseases with no currently available therapeutic intervention (e.g. Huntington’s disease) or for family planning.

In the United States the lack of a precise definition has produced a practical impediment to introduction of many “clinical” tests because, since 1974, by statute the FDA requires that all devices, including tests, must demonstrate “clinical utility” before they can be approved or cleared for entry to the US market. Without a clear and consistent definition of “clinical utility” in FDA regulations or guidelines, test developers have had to make a “best guess” at how to meet their interpretation of this requirement and hope for agreement with the FDA, often after the test has been fully analytically and clinically validated.

Developers and manufacturers of clinical tests, including genetic tests, should be responsible for establishing analytical and clinical validity. Information on clinical utility should be heavily dependent on medical and technical literature citations, be updated as information becomes available and be required in the product labelling. In any event, as long as the information is current and cited from reputable sources, i.e. peer-reviewed journals, it should not retard the availability of the test or device.

We suggest that the definition of “clinical utility” be expanded to incorporate the medical and the psychological needs of the patient: clinical utility refers to “the potential benefits in diagnosis, therapy, and the benefit of providing information about the propensity towards developing symptoms or a disease or for making decisions relevant to life choices, such as family planning”.

It should be noted that as more information becomes available, with regard to population studies, the clinical utility of a test is likely to change. The increase in such information *is expected and encouraged*. The performance of a particular test and what useful information it can provide will vary depending on which purpose it is used for and which population is being tested.

3 Working Group Session

3.1 Workshop participants were invited to join one of three working groups. Each group was asked to consider a series of specific questions related to genetic test evaluation, which are shown below:

1. What are the key questions?
2. What are the key components?
3. What are the main barriers to performing evaluation?
4. What is required to obtain the necessary information?
5. Which bodies or organisations should perform the evaluation?

3.2 Summary feedback: The key questions in genetic test evaluation are to:

- Provide clear accurate definitions of what genetic test is being evaluated, and adopt an agreed terminology (assay, condition, target population, setting)
- Provide clear accurate basic parameters including all available information about technical assay options, statistical parameters, target population, and test setting (for rare diseases clarity regarding lack of evidence is essential)
- Improve clarity about conclusions that are population specific
- Provide indications for which testing will be offered (confirmation of diagnosis – prenatal, newborn, adult - carrier testing, presymptomatic)
- Allow inclusion of the available evidence using an appropriate weighting of the evidence (all colours) including grey literature
- Ascertain if there is an intervention that flows from the test result
- Ascertain if there is there an alternative test available to address the clinical problem
- Ensure that an evaluation is carried out at a point in the life cycle of the test to permit effective policy decisions
- Incorporate tests for susceptibility testing – targeting of populations.
- Allow evaluation of tests that inform patient choices and have benefits to the family and carers of the patient

3.3 The summary feedback for the question “what are the key components of genetic test evaluation?” was that:

- The ACCE framework needs technical development but is useful for test evaluation and can usefully be extended to evaluations of tests for rare conditions
- The ACCE framework is not a complete methodology and needs development according to the purpose of the test
- Evaluations will identify gaps in knowledge
- An evaluation is valid at a single point in time but may require revision
- Available models lay out the core questions; however, a data template suitable for evaluation is required
- Clarity is needed on the quantity and quality of data required for evaluation
- A process is required that verifies the truth of claims and facilitates yes/no decisions about whether a test should be available in a publicly funded health system
- The process should include both pre- and post-market evaluation
- The evaluation process should be open to being put into a risk management framework

3.4 The summary feedback for the question “what are the main barriers in performing genetic test evaluation?” was that:

- There is frequently a lack of high-quality evidence
- There is a lack of funding for: large-scale studies, translational research, and evaluation of the evidence
- Gene patents may restrict access to service providers and quality and quantity of available data, and can the impede collaboration necessary for effective evaluation
- Clinicians are insufficiently informed about genetic testing and this may impede the collection of data for evaluation
- The term clinical utility is not clearly defined
- Roles and responsibilities of agencies involved in test evaluation are often poorly defined
- Any potential conflicts of interest amongst those supplying evidence and the evaluators should be identified and addressed
- Poor definition of semantics and terminology
- No platform or standard process in any OECD country for evaluation of genetic tests exists
- There is a need for public and professional education to encourage informed consumer questioning of genetic tests
- A better understanding of genetic test outcomes including potential harms is required

3.5 The summary feedback for the question “what is required to obtain the information necessary to perform genetic test evaluation?” was:

- Appropriate search strategies and critical evaluation of available literature
- A need for governments and responsible bodies to facilitate data collection
- Multi-centre (international) collaborations to aid in the collection and sharing of relevant data to inform the ongoing evaluation of genetic tests

3.6 The summary feedback for the question “which bodies or organisations should perform genetic test evaluation?” was:

- Evaluation requires an interdisciplinary approach and should be carried out by a competent medical and/or scientific body
- The evaluator should inform but be independent of the policy maker
- Different perspectives from purchasers, clinicians and patients should be taken into account
- The evaluation body should include all stakeholder views
- HTA organisations, although increasingly effective and collaborating internationally, need to clarify their role and facilitate data generation and analysis

4 Summary of Main Points from the Meeting and Considerations for Policy Action

- 4.1 The purpose of genetic test evaluation is to ensure that genetic tests provide a benefit or at least do no harm to an individual or the population as a whole.
- 4.2 Evidence presented during the workshop revealed that significant genetic test evaluation activity was being carried out in OECD member countries.
- 4.3 The meeting attendees agreed on the importance of genetic test evaluation to the development of genetics and clinical genetics services. It is also considered important in the development of health policy.
- 4.4 It was acknowledged that the infrastructure and processes for evaluation would be country and health system specific, as would the final judgement on clinical utility and its relation to the availability of a genetic test. The various national approaches to genetic test evaluation differed and each had its strengths and weaknesses. The different models of implementing genetic test evaluation included health technology assessment, government advisory groups, collaborative laboratory networks, public/government agencies and academic and government partnerships. The outputs of the evaluations performed also varied and included reports, guidelines, healthcare policy and direct provision of tests following approval.
- 4.5 There was a consensus amongst workshop participants on the domains of genetic test evaluation - disorder/context, analytical validity, clinical validity, clinical utility and ethical, legal and social issues. However, further work was required to obtain international consensus to establish or refine definition, terminologies, sub-sets and dimensions of these domains and the criteria by which they should be evaluated.
- 4.6 It was agreed that one of the main barriers to performing genetic test evaluations was the lack of evidence and data for clinical validity and clinical utility. In particular, there was limited information on clinical outcomes of testing. One approach to improve this would be to establish robust systems for the collection of post-implementation data. The benefit and harms of genetic testing need to be defined better and considered in evaluations. A need was also identified to develop quality standards for evidence and data.
- 4.7 It was agreed that infrastructure development and the creation of national and international networks to share data that could be used in genetic test evaluation are of high priority. This would provide a mechanism for making evidence internationally generalisable. It was suggested that the Health Technology Assessment system of evidence synthesis might be an appropriate model to promote such standards of evidence.

Considerations for policy action

It was agreed that there was scope for international action on genetic test evaluation. Consideration could include:

1. Developing an agreed framework for genetic test evaluation.
2. Establishing a formal international network approach to genetic test evaluation.
3. Considering what processes and infrastructure are necessary to develop the function of genetic test evaluation in their individual countries.
4. Establishing incentives and accountabilities for performing genetic test evaluation.
5. Developing consensus quality standards for data and evidence in order to allow thorough clinical validity and clinical utility analysis of genetic tests.
6. Developing consensus guidelines on genetic test evaluation accessible to the public and health professionals.

5 Acknowledgements

The project team would like thank the expert advisory group members, workshop speakers and participants for their enlightening and valuable contributions. We also would like to give special thanks to Dr Elettra Ronchi, Biotechnology Unit, OECD, who provided the core support to develop and present the workshop.

Finally, this workshop was only possible with the support and leadership of the OECD and we are grateful for this opportunity to contribute to the development of the area of genetic test evaluation.

6 Appendices

Appendix 1

Expert advisory group

An expert group comprising the following members developed the workshop programme:

Dr Ingeborg Blancquaert	Agence d'évaluation des technologies et des modes d'intervention en santé, Montréal, Canada
Dr Linda Bradley	National Office of Public Health Genomics, CDC, Atlanta, United States
Professor Angela Brand	German Center for Public Health Genomics (DZPHG), Bielefeld, Germany
Professor Wylie Burke	University of Washington, Seattle, United States
Dr Mark Kroese	Public Health Genetics Unit, Cambridge, United Kingdom

Appendix 2



DSTI/STP/BIO(2006)17
Unclassified

Unclassified

DSTI/STP/BIO(2006)17

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

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

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

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

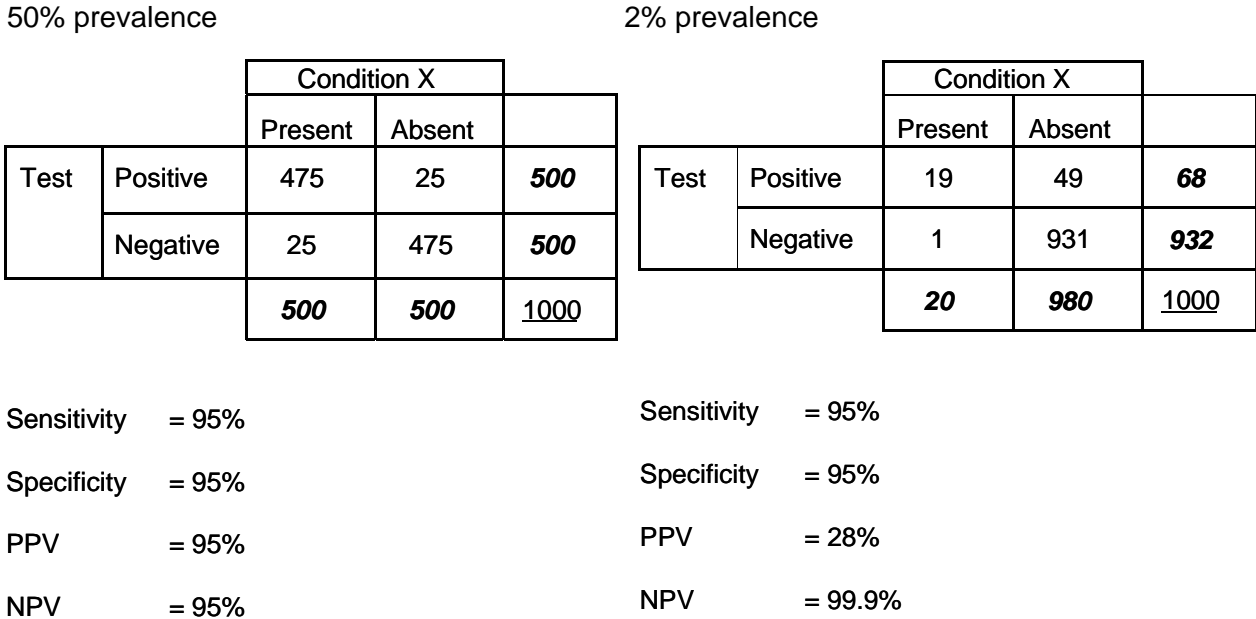
Sensitivity = $a/(a+c)$
 Specificity = $d/(b+d)$

PPV = $a/(a+b)$
 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



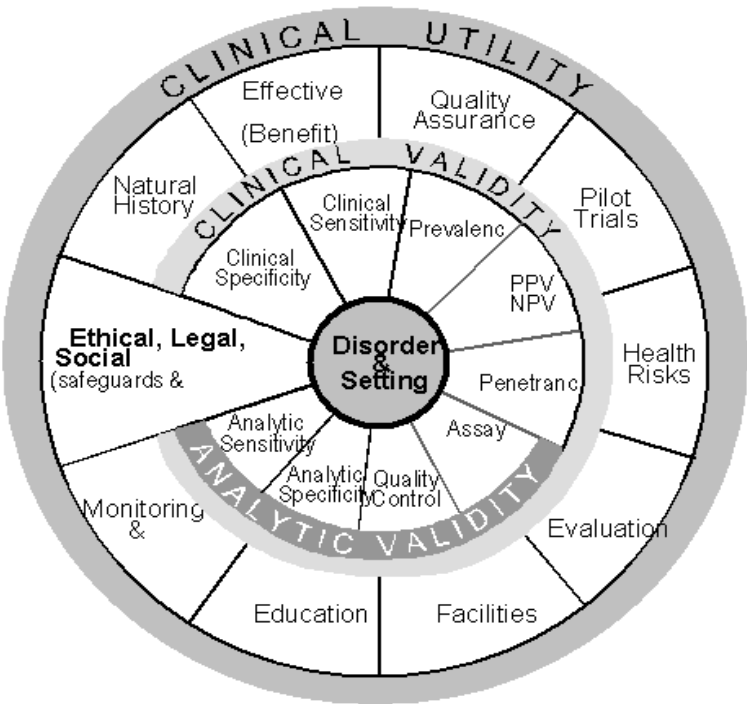
18. In order to establish the analytical and clinical validity of a genetic test, controls need to be included in the assessment. For analytical validity, negative controls are samples that are known not to have the mutation of interest; whilst for clinical validity the controls are individuals that 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 Framework²¹



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 quantitative 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:
- The test is not one of perfect analytical specificity.
 - The correct genotype has been identified but the phenotype is not present because of reduced penetrance.
 - 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:
- The test is not one of perfect analytical sensitivity.
 - Genes other than the one tested for are responsible for the disease (genetic heterogeneity).
 - Variants other than those tested for are responsible for the disease (allelic heterogeneity).
 - 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 emphasizes 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 policies 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.

Appendix A

Element	Specific Question
Disorder/ Setting	<ol style="list-style-type: none"> 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	<ol style="list-style-type: none"> 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	<ol style="list-style-type: none"> 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	<p>26. What is the natural history of the disorder?</p> <p>27. What is the impact of a positive (or negative) test on patient care?</p> <p>28. If applicable, are diagnostic tests available?</p> <p>29. Is there an effective remedy, acceptable action, or other measurable benefit?</p> <p>30. Is there general access to that remedy or action?</p> <p>31. Is the test being offered to a socially vulnerable population?</p> <p>32. What quality assurance measures are in place?</p> <p>33. What are the results of pilot trials?</p> <p>34. What health risks can be identified for follow-up testing and/or intervention?</p> <p>35. What are the financial costs associated with testing?</p> <p>36. What are the economic benefits associated with actions resulting from testing?</p> <p>37. What facilities/personnel are available or easily put in place?</p> <p>38. What educational materials have been developed and validated and which of these are available?</p> <p>39. Are there informed consent requirements?</p> <p>40. What methods exist for long term monitoring?</p> <p>41. What guidelines have been developed for evaluating program performance?</p>
ELSi	<p>42. What is known about stigmatization, discrimination, privacy/confidentiality and personal/family social issues?</p> <p>43. Are there legal issues regarding consent, ownership of data and/or samples, patents, licensing, proprietary testing, obligation to disclose, or reporting requirements?</p> <p>44. What safeguards have been described and are these safeguards in place and effective?</p>

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.

Appendix 3

Programme Expert Workshop “The evaluation of clinical validity and clinical utility of genetic tests”, 26 & 27 June 2006

DAY 1 - MONDAY, 26 JUNE 2006		
08:30 – 09:00	REGISTRATION	
09:00 –09:30	OPENING Welcome remarks Benedicte Callan (OECD)	
	INTRODUCTORY REMARKS: Rob Elles (<i>United Kingdom</i>) ; Elettra Ronchi (OECD)	
09:30 – 10:00	KEYNOTE ADDRESS	Introduction and Policy context <i>Ron Zimmern, United Kingdom</i>
10:00 –13:00	SESSION 1:	Current practice and developments in OECD member states Chair: Glenn Palomaki (US)
10:20 –10:40		<i>Peter Farndon, United Kingdom</i>
10:40 –11:10		<i>Glenn Palomaki, United States</i>
11:10 – 11:30		<i>Ulf Kristofferson, EuroGentest</i>
11:30 – 11:45	COFFEE BREAK	
11:45 – 12:05		<i>Cedric Carbonneil, France</i>
12:05– 12:20		<i>Gerardo Jiménez Sánchez, Mexico</i>
12:20 – 12:40		<i>Giuseppe Novelli and Domenica Taruscio, Italy</i>
12:40 – 13:00		Plenary Discussion
13:00 – 14:30	LUNCH	

14:30 – 18:00	SESSION 2:	Evaluation of clinical validity and clinical utility of genetic tests Chair: Angela Brand (Germany)
14:30 – 15:00		Challenges in Genetic Test Evaluation: Clinical Validity and Clinical Utility <i>Wylie Burke, United States</i>
15:00– 15:30		Testing for Rare Diseases : The UK experience <i>Fiona Stewart , United Kingdom</i>
15:30 – 16:00		Testing for Pharmacogenetics : The US Experience <i>Al Berg, United States</i>
16:00 -16:30	<i>COFFEE BREAK</i>	
16:30 – 17:00		Testing for BRCA: The Canadian Experience <i>Ingeborg Blancquaert , Canada</i>
17:00 - 18:00		Plenary Discussion
19:00 SOCIAL DINNER		
DAY 2 – TUESDAY, 27 JUNE 2006		
9:00 - 9:30	Patients Perspectives <i>Alastair Kent, United Kingdom</i>	
9:30 - 11:15	Working Group Parallel Sessions <i>Experts will be assigned to four groups and asked to review and discuss issues raised in earlier sessions. Each group will have a Moderator and a Rapporteur, who will be responsible for putting together a summary of main points and a short presentation at the end of the meeting.</i>	
11:15 - 11:30	COFFEE BREAK	
11:30 – 12:30	Working group reports	
12:30 – 14:00	LUNCH	

<p>14:00 – 16:00</p>	<p style="text-align: center;">Roundtable Debate Implications for Policy <i>Moderators: Elettra Ronchi (OECD)/ Ingeborg Blancquaert (Canada)</i></p> <p style="text-align: center;"><i>Discussants:</i> <i>Ronald Trent , Australia</i> <i>Glenn Palomaki, United States</i> <i>Evans Gareth/Jayne Spink , UK</i> <i>Ron Zimmern , UK</i> <i>Francois Thépot , France</i> <i>Dietmar Vybiral , Austria</i></p>	
<p>15:45 – 16:00</p>	<p style="text-align: center;">COFFEE BREAK</p>	
<p>16:00 – 16:30</p>		<p>Summary of Main Points from the Meeting and Recommendations for Policy Action <i>Mark Kroese, United Kingdom</i></p>
<p>16:30 - 17:00</p>		<p>Closing Remarks <i>Benedicte Callan (OECD); Rob Elles (United Kingdom)</i></p>

Appendix 4

Workshop Participants

Australia

Dr. Greg ASH
Acting Executive Director
National Health and Medical Research Council of Australia
GPO Box 1421
CANBERRA ACT
2601
Australia

Professor Ronald TRENT
Chair
Human Genetics Advisory Group
National Health and Medical Research Council of Australia
GPO Box 1421
CANBERRA ACT
2601
Australia

Austria

Mr. Gerald HÖFLER
Medical University of Graz
Institute for Pathology
Auenbrugger Platz 25
A-8036 Graz
Austria

Mr. Dietmar VYBIRAL
Federal Ministry for Health and Women
Biotechnology and Genetic Engineering
Radetzkystrasse 2
A-1030 VIENNA
Austria

Canada

Dr. Ingeborg BLANCQUAERT
Agence d'évaluation des technologies et des modes d'intervention
en santé
2021, avenue Union , bureau 1040,
MONTREAL
QUEBEC
Canada

Dr. Martin SOMERVILLE
University of Alberta
Department of Medical Genetics
8-26 Medical Sciences Bldg
Edmonton
Alberta T6G 2H7
Canada

Czech Republic

Professor Radim BRDICKA
National Ref. Lab. DNA
Institute of Hematology and Blood Transfusion
U. Nemocnice 1
12820 PRAGUE
Czech Republic

Denmark

Dr. Lars HANSEN
Novo Nordisk A/S
Krogshoejvej 53A
DK2 880 Bagsvaerd
Denmark

France

Mme Elisabeth BURG
AFSSAPS
143-147 BD Anatole
93285 Saint Denis
France

Dr. Cédric CARBONNEIL
Chef de projet
Haute Autorité de Santé
Service Evaluation des Actes Professionnels
2 avenue du Stade de France
93218 Saint Denis
France

Prof. François THEPOT
Agence de la biomédecine
1 Av. Du Stade de France
93212 cedex Saint Denis

Germany

Professor Angela BRAND
German Center for Public Health Genomics (DZPHG)
University of Applied Sciences Bielefeld
Kurt-Schumacher - Str.6
33615 Bielefeld
Germany

Mr. C. Micha NUBLING
Head, Section of "Molecular Pathology"
Paul Ehrlich Institute
63225 Langen
Germany

Mr. Tobias SCHULTE IN DEN BAUMEN
+49 521 106 7865
Public Health Genomics European Network (PHGEN)
Kurt Schumacher Str. 6
33615 Bielefeld
Germany

Italy

Dr. Giuseppe NOVELLI
Università Tor Vergata
Biopatologia e Diagnostica per Immagini
II Università Tor Vergata
I-00133 ROMA
Italy

Dr. Domenica TARUSCIO
Istituto Superiore di Sanità
Rare Diseases Unit
Viale Regina Elena 299
00161 ROMA
Italy

Mexico

Dr. Gerardo JIMENEZ SANCHEZ
Director General
The National Institute of Genomic Medicine
Mexican Ministry of Health
Mexico

Netherlands

Professor Egbertt BAKKER
Laboratory for Diagnostic Genome Analyses (LDGA)
Centre for Human and Clinical Genetics
S-6-P P.O. Box 9600
2300 RC Leiden
Netherlands

Norway

Dr. Gunnar HOUGE
Center for Medical Genetics and Molecular Medicine
Haukeland University Hospital
5021 Bergen
Norway

Dr. Lovise MAEHLE
Radiumhospitalet - Rikshospitalet HF Street
0310 Oslo
Norway

Portugal

Dr. Jorge SEQUEIROS
UnIGeNe - IBMC, Univ. Porto
R. Campo Alegre, 823
4150-180 PORTO
Portugal

Spain

Dr. Javier GARCIA-PLANELL
Responsable de Unidad de Genética Médica
Sistemas Genomicos
Paterna
Ronda G. Marconi, 6
46980 Valencia
Spain

Ms. Cecilia MARTÍN BOURGON
Conseiller à la Direction des Affaires
Ministère de la Santé et de la Consommation
Institut de Santé Carlos III
MADRID
Spain

Sweden

Dr. Ulf KRISTOFFERSSON
Associate Professor
University Hospital
Department of Clinical Genetics
Genetiska Kliniken
SE-22185 LUND
Sweden

United Kingdom

Dr. Robert ELLES
Head
National Genetics Reference Laboratory

Hathersage Road
MANCHESTER
M13 OJH
United Kingdom

Professor Peter FARNDON

West Midlands Regional Clinical Genetics Service
Birmingham Women's Hospital
Edgbaston
Metchley Park Road
Birmingham
B15 2TG
United Kingdom

Dr. Alastair KENT

Director
Genetic Interest Group
4D Leroy House,
436 Essex Road
LONDON
N1 3QP
United Kingdom

Dr. Mark KROESE

Public Health Genetics Unit
Strangeways Research Laboratories
Cambridge
CB1 8RN
United Kingdom

Ms. Diana PAINE

Department of Health
133-155 Waterloo Road
London
SE1 8UG
United Kingdom

Dr. Fiona STEWART

City Hospital
Department of Medical Genetics
Belfast
BT9 7AB
United Kingdom

United States

Dr. Wylie BURKE

University of Washington
1959 NE Pacific
Seattle
Washington 98195-7120
United States

Mr. Fred LASKY

Genzyme
500 Kendall Street
Cambridge MA
02142
United States

EC

Dr. Dolores IBARRETA
Institute for Prospective Technological Studies
European Commission
Edf. EXPO - C/. Inca Garcilaso, s/n
E 41092 Seville
Spain

Lithuania

Mr. Vaidutis KUCINSKAS
Vilnius University
Santariskiu 2
08661 Vilnius
Lithuania

OECD

Ms. Benedicte CALLAN
Administrator
OECD
STI
2, rue André Pascal
75016 Paris
France

Ms. Elettra RONCHI
OECD
STI
2, rue André Pascal
75016 Paris
France

Other

Dr. Ron ZIMMERN
Public Health Genetics Unit
Worts Causeway
Strangeways Research Laboratory
Cambridge
CB1 8RN
United Kingdom

Dr. Joanne WHITTAKER
Regional Genetics Laboratories
Maris Lane
Trumpington
Cambridge
CB2 2FF
United Kingdom

Dr. Alfred BERG
Professor and Chair
University of Washington
Department of Family Medicine
United States