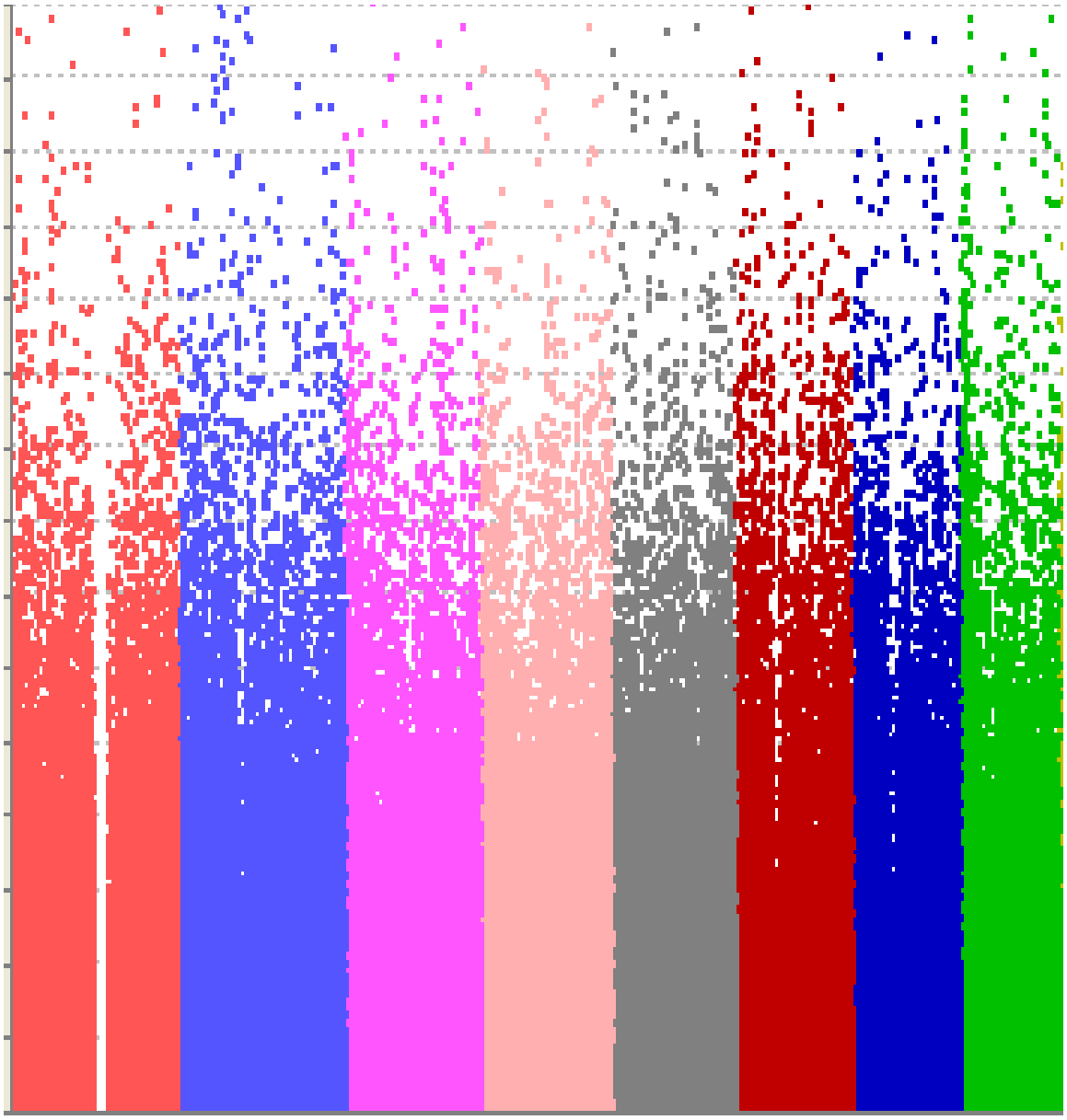
EVIDENCE AND EVALUATION: BUILDING PUBLIC TRUST IN GENETIC TESTS FOR COMMON DISEASES

Research Report



Professor David Melzer, Peninsula Medical School, Exeter University, UK
Stuart Hogarth, Department of Public Health and Primary Care, University of Cambridge
Dr Kathy Liddell, Faculty of Law, University of Cambridge
Professor Tom Ling, RAND Europe
Dr Simon Sanderson, PHG Foundation, Cambridge
Dr Ron Zimmern, PHG Foundation, Cambridge

February 2008

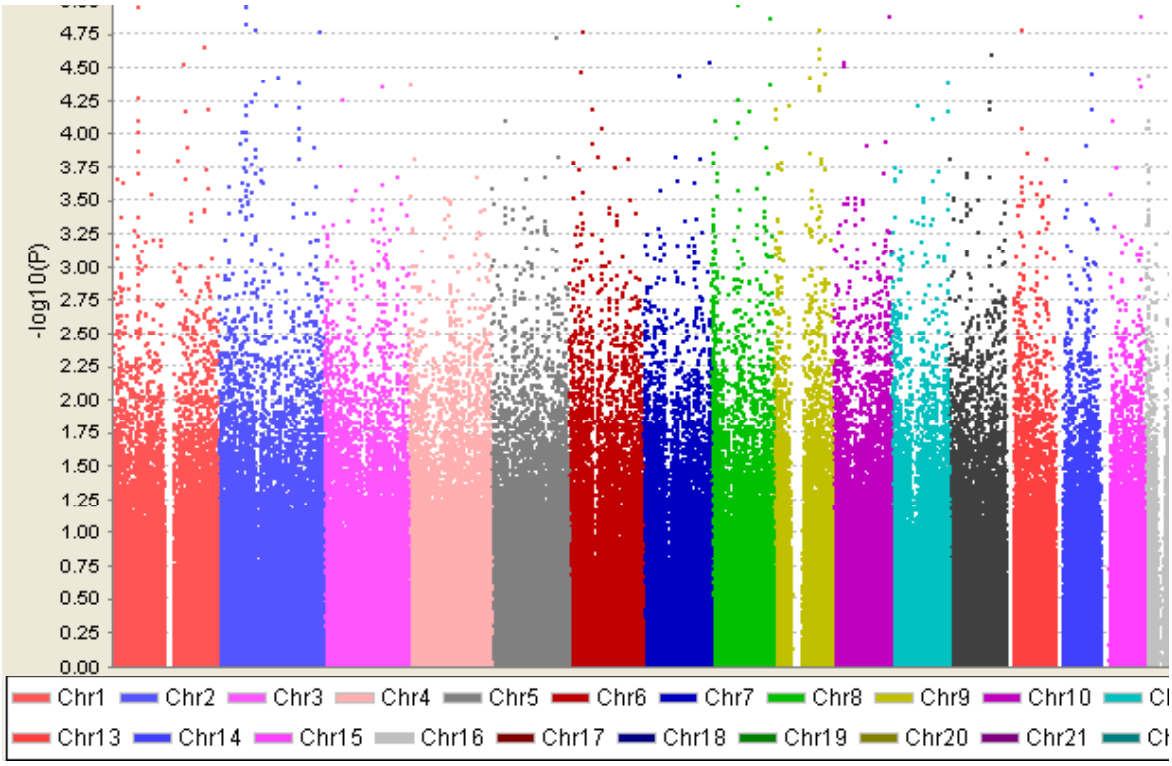
phg foundation
making science work for health

PENINSULA MEDICAL SCHOOL
UNIVERSITIES OF EXETER & PLYMOUTH

Cover illustration

Detail from genome wide association plot of variants by physical position on each chromosome against statistical chance of association with a key inflammation marker, Interleukin 6 levels in blood

See Rafiq S, et al. Genes & Immunity 2007 Oct;8(7):552-9. A common variant of the interleukin 6 receptor (IL-6r) gene increases IL-6r and IL-6 levels, without other inflammatory effects.



EVIDENCE AND EVALUATION: BUILDING PUBLIC TRUST IN GENETIC TESTS FOR COMMON DISEASES

Introduction

The new genetics has moved with extraordinary speed over the last decade. Since the sequencing of the human genome, over ten million DNA variants between people have been characterised. In the past most genetic tests were for single-gene disorders in which having the gene variant was synonymous with having the target condition. Recent genome wide studies have identified a rapidly expanding set of robustly proven associations between inherited gene variants and common diseases. There have been exciting findings for diseases including myocardial infarction, diabetes, age-related macular degeneration, asthma and several auto-immune conditions. However, while some of the emerging markers have moderate or larger effect sizes, suggesting a reasonably strong association between the marker and the disease, many indicate only small degrees of risk. This picture is further complicated by many earlier exploratory analyses which suggested associations between hundreds of variants and disease states, most of which have subsequently proved weak or absent altogether.

Relatively cheap laboratory assays can now identify genetic variants in individual samples and in large collections. Innovation and development in genetic testing increasingly lie in establishing the clinical significance of assay results: establishing what a positive or negative result actually means, and what to do about the result in individual patients. Many of the emerging genetic assays appear to have the potential to become useful clinical tests for estimating future risk of common diseases, particularly in diagnosis, estimating risk, guiding treatment, and determining prognosis. However, formal research is needed to identify the clinical uses to which these markers could be validly and usefully applied: in most cases the current scientific evidence statistically linking variants to complex disease is not sufficient to support clinical use.

A consensus is emerging about the need to distinguish the laboratory analysis (or 'assay') determining the presence of a particular genetic variant, from the clinical test. The clinical test should be seen as the application of the assay for a particular purpose and in a particular population. The evaluation of a laboratory analysis or assay is technically straightforward, allowing broadly applicable standards to be established. The evaluation of a clinical test is more complex and inherently less standardised, as the significance of a result is critically dependent on context: each clinical test application has to be evaluated on a case-by-case basis. The information needed to enable a test to be properly evaluated includes the:

- specific purpose of the test (eg, diagnostic, predictive, prognostic etc.)
- the disease or disorder at issue
- nature of the genetic variants being identified, and the **analytic validity** of the analysis (i.e. how well the genetic variants are ascertained)
- target patient group or **population**

- **the clinical validity** (e.g. the evidence underpinning the disease prediction)

- **the clinical utility** (evidence that the test result changes clinical care in a useful way and leads to improved health)

This information would enable the meaning of the results for an individual patient to be interpreted and major areas of uncertainty clarified.

If tests provided to consumers, patients or doctors perform differently to what is claimed or implied in promotional literature or test reports, then the results have the potential to seriously mislead clinical decision-making. In areas such as pharmacogenetics, misleading performance can result in direct harm. In other areas the main effects of false positive and negative results may be a cascade of unnecessary follow-up testing and treatments.

Much of the harm of poorly evaluated clinical tests goes unnoticed, as errors play contributory rather than direct roles in medical misadventures.

Rarely does test performance, however poor, attract political or media attention sufficient to elevate clinical evaluation of tests onto the policymakers' priority list.

The scope of the research and the report

Over the past 15 years, experts have predicted that genetic testing would play a greater role in medicine; but at the same time concerns were also expressed that some genetic tests were entering clinical practice prematurely and without adequate evaluation. As one senior diagnostics industry figure put it:

[There has been] a noticeable lack of consensus within the genetics community about exactly when a test for a new marker was sufficiently validated for it to enter into clinical service. Some labs rushed to provide testing after the first publication, while others waited until the result had been replicated in multiple studies or multiple ethnic groups[1].

This independent academic research report focuses on the factors influencing how new genetic tests for common disease susceptibility enter routine clinical practice, and at the need for appropriate clinical evaluation. It looks in detail at:

- the generation of evidence on test performance, including incentives for test developers to undertake clinical studies

- the evaluation of evidence, including the roles of regulators, reimbursers and professionals

- the dissemination of evidence on each test

The central concern is that consumers, patients and doctors should have sufficient information about a test to decide whether to use it and how to interpret the result.

This report reflects the results of a research project involving over 80 interviews with opinion leaders from all the key parties involved in the clinical use of genetic tests for common conditions. In this report we have tried to identify both the areas of agreement

[1] Winn-Deen, Emily S. 'Fulfilling the promise of personalized medicine' IVD Technology, November/December 2003

and of disagreement. We have also identified core principals and potential ways forward, representing our own synthesis of the material examined.

Key findings

The great majority of clinicians and patients are unable to either keep track of the burgeoning scientific literature on genetic assays, or learn in the course of everyday clinical experience the utility of the tests in specific clinical situations. This is especially true for tests that indicate susceptibility to disease into which most of the emerging common disease variants are likely to be incorporated. A proper clinical evaluation of genetic assays and tests was seen by those interviewed as being in the interests of patients, doctors, labs and device manufacturers and to aid commercial innovation and success in the market place.

There was therefore near universal support for improving the generation of good clinical evidence on tests and for this evidence to be made easily available to doctors, patients and consumers.

The pathway from bench to bedside is controlled by a series of gatekeepers or points of control. Broadly speaking these operate at three levels: statutory controls, resource allocation and clinical governance.[2] A new genetic test might therefore be regulated at:

- the first level by standards set by a statutory licensing body, such as the Food and Drug Administration (FDA) in the USA or the responsible bodies under the EU directive on in-vitro diagnostics

- the second level by the requirements established by a purchaser, commissioner or reimburser of services, such as the UK's National Health Service

- at the third level by the rules and guidelines set by professional bodies, healthcare organisations and other groups, which set standards in the practice of medicine

None of the gatekeepers or points of control is sufficient on its own to satisfy the need for good clinical evaluation and accessible evidence

Incentives and resources to generate clinical evidence

One of the key requirements for achieving better clinical evaluation of genetic tests is to develop a framework of incentives and resources that will allow the clinical data to be generated before introducing a test into routine use. The resources required will often be modest in comparison with the cost of clinical trials for new medicines, although the cost of such studies might still be too great to be carried by the biotechnology industry on their own. There has been a marked growth in the use the use of patents to protect new genetic discoveries and to generate a stream of resources, but questions remain about the effectiveness of the patent system in protecting the intellectual property of diagnostics. Some companies have changed their business models from providing assay infrastructure to providing test services to capitalise on the IP regime, but there remains considerable doubt as to whether these changes are sufficient on their own to encourage innovation and evaluation within the commercial sector.

[2] Burke W. and Zimmern R.L. (2004) 'Ensuring the appropriate use of genetic tests' Nat Rev Genet. Dec;5(12):955-9

At present there is little consensus as to who should be responsible for providing the resources or the infrastructure to allow such data to be generated. It appears that no single party (government, the health system or the commercial sector) has either the responsibility or the resources to undertake adequate clinical studies of the validity and utility of test applications. Many of the participants of the project agreed that public funding should be used to address this issue, probably in partnership with the private sector.

Whatever actions are taken, good clinical evaluation will require more than the demonstration of statistical association between genetic variant and disease. Clinical studies will be needed to determine test characteristics. These may be carried out by test providers themselves, or health systems and academic institutions. The resources would have to come either from industry through profits from improved IP mechanisms, or through explicit policy mechanisms that enable public funds to be used for the collection of the necessary data. The two approaches are not mutually exclusive and small incremental steps to developing both will probably be most effective in practice.

Evaluating evidence: statutory controls

Statutory controls are used to provide an initial assessment of the accuracy of the assay (its analytic validity) and depending on jurisdiction its clinical validity also. A statutory mechanism to ensure truth-in-labelling (that the claims made by test providers are consistent with the available data on their test) is in place in most developed country markets. This was seen by many to be the key to proper regulation. It was necessary to ensure transparency of information, and to encourage the placing of evidence and data for the clinical validity and utility of tests in the public domain so that all stakeholders – reimbursers, physicians and patients – might have access to the data. It was not that tests without adequate evidence should necessarily be refused entry to the market (unless issues of safety were involved), but that the evidence as to the test's validity or utility should be made available for all to see.

The 'biological plausibility' of a clinical test (that is prima facie evidence of a relationship between the genetic variant under discussion and a particular disorder) was not deemed on its own to constitute adequate evidence of clinical validity, except perhaps for high penetrance genetic tests for inherited disorders. Some regulators, notably the US FDA, can require formal evidence of clinical validity and utility, depending on the tests' intended use, but the standard of evidence required is unclear, and there are examples where strong evidence of statistical association between genetic variant and disease appeared on its own to fulfil the FDA's requirements.

The development of formal regulatory systems in the USA, Europe and other developed countries has taken a variety of routes, but has inevitably faced many common challenges. The US Food & Drug Administration has a pre-market approval role for test devices, which includes examination of clinical data, although standards for these data are set on a case-by-case basis. In Europe a system of self-certification with supporting documentation operates for genetic tests: however, these product dossiers are currently kept secret from consumers, patients and doctors. The dossiers may be called in by regulatory agencies if concerns arise, although secrecy impedes the professional or public identification of possible problems. The clinical content of these self-certification dossiers is thought to be low, as tests that were rejected by the US-FDA as lacking any evidence of clinical validity have claimed "CE" marking in Europe.

The systems of statutory regulation in Europe and the USA were both felt to have major

areas of weakness. Neither regulatory system was felt to currently deliver the clinical evidence that clinicians and patients need. The exclusion of the great majority of laboratory-developed tests from FDA scrutiny in the USA introduces a perverse incentive to market tests as lab services rather than (regulated) test devices. The European system is being interpreted as exempting public sector laboratories and so-called 'lifestyle' genetic testing, so both areas will need specific measures to ensure the clinical quality of testing.

The participants in the project were by and large unanimous about the need for a greater degree of clinical evaluation, but its exact form, the extent to which it should be part of the formal statutory regulatory process, and the responsibility for funding such studies were all matters that generated a range of views between, and within, different groups of stakeholders.

Developing the statutory role

There was general agreement about the principles that should govern the various aspects of the statutory role. Any regulatory system for genetic tests should be transparent, consistent and comprehensive. It must:

- safeguard patients and consumers from harm
- must provide doctors, patients and consumers with relevant information (and avoid gross information imbalances between seller and buyer which could lead to market distortions)
- balance the need to deter inappropriate use of tests of unproven value against the need to encourage innovation and ensure rapid access to useful new technologies which could improve health
- provide test developers with a clear understanding of what was expected of them – the standards that would be applied and the processes for gaining regulatory approval
- achieve common standards across all genetic tests, even where there may be multiple pathways to those common standards

The current shortcomings of statutory regulation were felt to derive from a number of fundamental difficulties arising from the way diagnostics providers and health systems operate:

- although few laboratories currently provide genetic testing for common conditions, very large numbers would have the technical capability to do so in the future, raising serious issues about regulatory logistics
- clinical evidence about the validity and utility of a test is complex and dependent on the details of each application: universal standards for the evidence required could only be defined in broad terms and a case-by-case approach would be needed to decide the merits of each test for each target population and each clinical purpose, again posing logistical challenges
- some laboratories see themselves as only undertaking assays and make no claims for the clinical validity of these: although current commercial services offering

genetic testing to the public and doctors all appear to make clinical claims, it may be that labs in the future will choose not to do so

- in the past there has been a lack of political will to improve the clinical evaluation of tests, perhaps because the harm of misleading tests tends to be hidden

Given these difficulties, there was disagreement about the appropriate scope of formal regulation, especially in relation to laboratories working as support services to doctors. Neverthless, in the light of the large numbers of new tests now emerging (including from genetics), there is an international trend towards regulating some or all laboratory-developed tests as medical devices: this is now established in European and Australian oversight arrangements and developing in the United States.

Improved regulation of testing services making clinical claims, especially those marketing directly to the public, did appear to have widespread (although not universal) support. A minority of stakeholders feared a burdensome regime which would deny access to lower risk tests that consumers wanted.

Risk classification

A key step in statutory regulation is to assign individual tests to a risk level, distinguishing those that pose more serious dangers and therefore require greater scrutiny. The criteria for risk classification (well-established in USA regulation) include:

- novelty (the lack of an already approved test that is substantially similar)

- the lack of supporting sources of data (where decision making is wholly dependent on the test result)

- the impact of the test on clinical management of the patient (if the test result triggers potentially hazardous surgery or drug treatment)

- the public health implications of tests (for screening programmes in which small inaccuracies can lead to very large numbers of misleading results)

In Canada, Australia and the US, genetic tests have been placed in higher risk categories for enhanced regulatory oversight, ensuring that they are subject to pre-market review. However, virtually all genetic tests in the EU IVD Directive are exempt from pre-market scrutiny because they are currently placed in the lowest risk grouping (even those linked to the prescribing of potentially dangerous drugs or sold direct to the public). Given the nature of recently discovered markers and potential applications, we anticipate that some would be classified into at least moderate risk categories.

It was felt by most that some form of risk classification will provide the way forward for the statutory regulation of genetic tests. There was less consensus on whether a strict classification system could be embodied in a set of rules. We believe that flexible risk classification principles should be adopted, to be applied or overseen by an independent scientific committee with patient representation.

‘Light touch’ pre-market review

Another solution supported by many stakeholders was an emphasis on a light touch approach to premarket review. At its most ambitious regulators will use premarket review to set out in detail the types of clinical studies they require a test developer to perform for a specific submission. At its most modest, premarket review will focus on ensuring truth-in-labelling, using more general guidance to indicate the types of evidence which will be acceptable. Our research showed strong support for an approach focused on using premarket review to ensure truth-in-labeling (and truth-in-promotion) as a minimal approach to premarket review.[3]

This approach presents a minimal evidence requirement since it is possible for a test developer to rely solely on the existing scientific literature rather than conducting costly clinical studies of their own (providing of course that the literature supports the test developer’s intended use). Linked to this approach to pre-market review was the broadly-supported view that data on analytic and clinical validity are minimum data requirements for this stage of evaluation. It was generally accepted that it is both unrealistic to ask statutory regulators to evaluate the clinical utility of tests or their ethical, legal and social implications, and probably constitutes too high a barrier to market entry.

“In-house” developed tests

A further gap exists in the US and Canada where tests developed in-house by laboratories are generally not subject to the same pre-market review procedures as tests developed as kits. This anomaly has been addressed for private laboratories in Europe and Australia and there is widespread support for a more consistent approach in the US. IVD manufacturers in the US shared other stakeholders’ concerns about the lack of a level playing field between laboratory-developed tests (LDTs) and test kits. Recent guidance from the FDA indicates that the Agency is now beginning to address this, currently in the area of test results based on computer algorithms. Public sector labs in Europe are, however, exempt from the EU IVD directive, although alternative mechanisms for ensuring the clinical quality of in-house assays from these labs are available through professionally driven quality assessment schemes.

The European challenge

The current European system of regulation could provide a workable framework for the future if certain basic flaws are addressed. An independent scientific committee should be responsible for determining the risk category for new genetic tests, and the current system of this being decided by civil servants through a political process should end. For each test, a clear and publicly available statement should be made covering the basic clinical information on each test. If clinical claims are made the evidence supporting the test should be set out in a form easily available to doctors and patients (preferably in a structured summary on the internet, thus reducing the costs for consumers and regulators overseeing the test market). The procedure for reporting concerns about an assay or test to the competent regulatory authority should be publicised to consumers, patients and doctors. Regulators should ensure that systems for monitoring the genetic testing market and responding appropriately to concerns are in place.

[3] Also of relevance may be consumer law instruments for dealing with the claims made for products and services

The US challenge

The current US system was considered to have weaknesses in the area of laboratory provided tests, with the traditional lack of FDA oversight of this sector resulting in a perverse incentive to offer new tests as largely unregulated services rather than regulated devices. The FDA has identified some laboratory provided tests which it is now subjecting to pre-market review. A broader and more consistent approach to risk classification is required to identify other categories of test which require pre-market review. For lower risk tests, a move to a reactive system, combined with a duty to publish structured clinical information, could also provide one way forward in the USA context.

Health system and professional action

The health system route to clinical evaluation has developed to some extent, for example through the use of gene dossiers within UK Genetic Testing Network (UKGTN). However, such approaches are generally limited to reviewing the often scant available clinical data. The experience of the UKGTN is confined mainly to high penetrance tests for inherited disorders, where detailed evaluation is perhaps less important. It is nevertheless clear that the research infrastructure for test evaluation will have to be further developed and supported, and that the failure to have such an infrastructure should be a key concern of policy makers.

As noted above, much work is will be needed to improve the general genetic education of health professionals, and to provide access to relevant information on individual tests. Professional bodies and health care providers should remind professionals that using tests in routine practice without good clinical evaluation is not compatible with good clinical practice. Reimbursers should support the development of new tests in clinical research contexts but only pay for routine tests that have been adequately evaluated.

Conclusion

There is a consensus that genetic tests for common conditions should not enter routine clinical practice without the evidence to support their use. Its generation will require resources and infrastructure. However, neither test makers nor health systems have a clear responsibility to put these mechanisms in place. Current European device regulations place genetic tests in the lowest risk category; little clinical data is required in documentation and what data that exist are kept secret from doctors and patients. Current health system responses are generally limited to review of often limited existing data. Policy action is needed to develop the roles of both test provider and health systems in clinical evaluation. Failure to improve clinical evaluation would undermine the development of personalised medicine in the 21st century, and lead to a new generation of medical technology of unclear clinical value.

Specific Recommendations

1. The scope for testing genetic variants associated with common conditions has recently greatly expanded, with genome wide studies reporting many variants. A review of the effectiveness of current oversight arrangements for these tests is now needed.

2. In our interviews and focus groups with stakeholders, there was agreement that genetic markers should not be used routinely in clinical practice without the evidence to support their intended use. This requires that evidence of the clinical validity and utility of a test be generated, and made available in an open and transparent manner. Where little evidence exists or where the evidence is confined to statistical links between the marker and the disease, this should be explicitly stated. The principle that the evidence base should be placed in the public domain and available to all should underpin policy action.

3. The systems of statutory regulation in Europe and the USA differ markedly in their approach, but both were felt to have areas of weakness. Neither was felt to currently deliver the clinical evidence that clinicians and patients need. The shortcomings of statutory regulation were felt to derive from a number of fundamental difficulties arising from the way diagnostics providers and health systems operate. Given these difficulties there is a lack of consensus about the appropriate scope of formal regulation, especially in relation to laboratories working as support services to doctors. However, policy makers should recognise that there are many areas where action is strongly supported by most stakeholders, especially where test providers make clinical claims or where they market tests directly to the public.

4. In Europe the current system of regulation was felt to provide a workable framework for the future, provided certain basic flaws are addressed including:

- the appointment of an independent scientific committee responsible for determining the risk category for new genetic tests

- a requirement for a clear public statement by test providers on what genetic markers are being tested and whether clinical claims are being made. If clinical claims are made these and the clinical evidence supporting them should be set out in a form easily available to doctors and patients (preferably in a structured summary on the internet)

- the procedure for reporting concerns to the competent authority about individual assays / tests should be publicised, and regulators should set up systems to encourage professional and public awareness of this process and to investigate and intervene where concerns justify it

5. The USA system of regulation has specific challenges and opportunities, but it may be that a regulatory approach to laboratory developed tests that emphasizes public disclosure of evidence, a more consistent use of risk classification to identify those laboratory derived tests which need pre-market review, and reactive regulation for lower risk tests may provide a way forward.

6. Tests of medical relevance sold direct to the public should be treated as having made clinical claims, and requirements to publicly report supporting clinical evidence should be enforced.

7. No single party has the responsibility and resources to undertake adequate clinical evaluation of the validity and utility of test applications: public funding alone or in partnership with the private sector will be needed to support clinical studies to realise the potential to improve health by the application of the new genetic markers to clinical practice.

8. Professional bodies and health providers should remind practitioners that using poorly evaluated tests is not compatible with sound clinical practice.

Acknowledgements

This report has been developed as part of a project funded by the Wellcome Trust under its Bioethics funding programme. The Trust played no part in the writing of this report. We are grateful to the Trust for its support. We would also like to thank the project administrator, Therese Williams, for her assistance with the work undertaken for this report.

Disclosure statement

None of the authors feel they have conflicts of interest relating to this work and none have relevant commercial links.

Other publications from the project

Papers

'The new common disease genetic tests: new insights, old concerns' British Medical Journal forthcoming March 2008 D Melzer, S Hogarth, K Liddell, T Ling S Sanderson and R Zimmern

'Closing the gaps – enhancing the regulation of genetic tests using responsive regulation' Food and Drug Law Journal November 2007 S Hogarth, K Liddell, T Ling, S Sanderson, R Zimmern and D Melzer

'The regulation of direct-to-consumer genetic tests' Annual Review of Genomics and Human Genetics Forthcoming 2008 S Hogarth, G Javitt and D Melzer

Briefings

The IVD Directive and genetic testing: problems and proposals. A briefing presented to the 20th meeting of Competent Authorities, Lisbon, July 2007S Hogarth and D Melzer (Cambridge University, July 2007)

Regulating pharmacogenomics – an overview of developments in various countries and industry response to regulatory initiatives. A report for Health Canada S Hogarth, K Liddell, T Ling, D Melzer and R Zimmern (Cambridge University, March 2006)

The regulation of commercial genetic testing services – a briefing for the Human Genetics Commission S Hogarth, D Melzer and R Zimmern (Cambridge University, August 2005)

For copies of this report please contact:

PHG Foundation
Strangeways Laboratory
Worts Causeway
CAMBRIDGE
CB1 8RN

www.phgfoundation.org

Professor David Melzer
Peninsula Medical School
Exeter University
Barrack Road
Exeter
EX2 5DW

email: david.melzer@pms.ac.uk

www.pms.ac.uk