



*Professor Peter Furness
Vice President and
President-elect*

Evaluating new diagnostic tests

Professor Furness reports back from a meeting co-hosted by The Royal College of Pathologists and the PHG Foundation on 14–15 January 2008, entitled 'The evaluation of diagnostic laboratory tests and complex biomarkers'.

How do you obtain an authoritative opinion on whether or not a diagnostic test should be used? You might be tempted to answer: 'Ask a pathologist'. But what if the diagnostic test is new, expensive and based on limited scientific evidence, and you are a Clinical Director or NHS Commissioner who has to decide whether or not to provide a service?

There are obvious parallels with decisions about whether or not the health service should provide new drugs. But new drugs are subject to a sophisticated regulatory framework, demanding clinical trials and formal evaluation processes such as those performed by the National Institute for Health and Clinical Excellence (NICE).

Evaluating new diagnostic tests is difficult. There are several phases.

1. Analytical validity

Does it measure what it claims to measure, with sufficient reliability and accuracy? This is often the least difficult part of the question, but is the only one addressed by the current regulatory system under the EC's In Vitro Diagnostics Directive.

2. Clinical validity

This can be expressed as the sensitivity and specificity of a new test, and its positive and negative predictive value. But these measurements are hugely dependent on the population under study. A test that is useful for patients with a specific complaint in secondary healthcare setting might have a completely different (and unacceptable) predictive value if applied to the general population as a screening test.

3. Clinical utility

This is the most difficult and the most rarely evaluated. Indeed it is often ignored. But it is crucial. Does the test actually generate patient benefit? There is no point in using a laboratory test where the benefit does not outweigh the (financial and therapeutic) costs. This can occur in a variety of ways, even if a tests diagnostic validity is excellent. For example, the condition might be diagnosed with adequate confidence by clinical history and examination, or a better test might exist.

4. The ethical, legal and social implications of the test

This is particularly likely to be important in the context of genetic testing.

Furthermore, unlike drugs, tests can be used in many ways; not just diagnosis, but also excluding diagnoses, guiding subsequent investigation, monitoring disease activity, monitoring treatment, population screening and others. Some might add 'Keeping the patient (or the doctor?) happy'.

This meeting discussed all these difficulties in considerable detail. There are several national organisations (including NICE) that have an interest in evaluating new diagnostic tests, but they all have other responsibilities, they all cover a limited part of the field and they evaluate only a small number of 'big impact' new investigations, often concentrating on new commercial products.

Evaluation is almost always bedevilled by a lack of evidence and the fact that most of the published evidence relates only to analytical and diagnostic validity under specific conditions, in specific popu-

Making Sense of Testing:
A guide to why scans and health tests for well people aren't always a good idea



Box 1:
Recommendations of
the meeting

1	A new body should be established to ensure the evaluation of laboratory diagnostic tests and the creation of a database of new and existing laboratory tests.
2	This body might be established <i>de novo</i> along the lines of the UK Genetic Testing Network, or the responsibility could be placed with existing professional societies such as The Royal College of Pathologists, the Association of Clinical Biochemistry or the Academy of Royal Colleges.
3	A publically available database of existing and new diagnostic laboratory tests should be set up containing evidence, or explicitly the lack of it, for the validity and utility of clinical laboratory tests.
4	Where a test evaluation has already been carried out and published by an appropriate agency it should be linked to the database.
5	Where evidence is missing for existing tests, particularly evidence of clinical validity and utility, consideration should be given to funding the necessary studies.
6	Policy makers and all stakeholders should be encouraged to address issues around funding and gathering the necessary evidence for the clinical evaluation of new and complex biomarkers, and should consider the establishment of private-public partnerships to increase industry involvement.
7	An independent expert body should be responsible for the evaluation of the evidence for test performance and making recommendations about clinical use.
8	Commissioners and health care professionals should be encouraged to use only those tests where sufficient evidence of clinical performance exists.
9	Statutory regulators should be empowered to require that evidence (or lack of) relating to test performance be placed in the public domain.
10	A more responsive and proportionate risk assessment during pre-market approval is needed to ensure patient safety.

lations. This lack of evidence is in part due to the absence of a proper regulatory framework; commercial companies bringing a diagnostic test to the market just do not face the demands placed upon pharmaceutical companies to demonstrate patient benefit. But even if they did, the profits to be gained would not justify the huge research budgets of the pharmaceutical companies, so the problem of insufficient evidence would persist.

So what's the solution?

A potential approach to this problem was described in the form of the 'gene dossier' system of evaluation developed by the UK Genetic Testing Network. This is a system of accumulating evidence and making a decision on whether a genetic test should be offered to NHS patients; much less rigorous than a full NICE evaluation, but nevertheless proportionate and adequate for the task. As yet, it has only been applied to monogenetic diseases.

The meeting came to a number of unanimous recommendations, listed in Box 1. Its full report is available (as a pdf file) at www.phgfoundation.org/file/3998/.¹ The solution is obvious but will be difficult to achieve. In brief, what is needed is a single national database of laboratory investigations, each with a summary of the circumstances under which the use of the test is (and is not!) valid. This

would be an equivalent to the British National Formulary, but would be directed to laboratory investigation rather than drugs. The available evidence for clinical validity and utility of each test should be summarised, in relation to situations where it is thought to be useful, including acknowledgement of the inadequacy of evidence. Having established this facility for current 'established' tests, a 'horizon scanning' function would be needed to identify and then evaluate possible new developments, for inclusion in the database. This would coordinate the results of several organisations that already have a horizon-scanning function, but should have considerably greater input from pathologists and relevant professional bodies. A mechanism for conducting evaluations would be needed – or, more likely, several proportionate mechanisms, ranging from a full NICE evaluation for developments of major national importance, to a much simpler evaluation for less costly tests that might be relevant only to a few patients. And there would be a need to stimulate research in those areas where rational decision making is blocked by lack of evidence. This resource would be of benefit to disparate users (Box 2). It would also encourage those who develop diagnostic devices to put more resources into research providing evidence of patient benefit, perhaps thereby spending less on marketing.

Box 2: Beneficiaries of a 'British National Formulary for diagnostic tests'

Clinicians needing to know whether the use of a test – or a demand for its implementation – is justified.
Pathologists needing support in deciding whether to develop or, importantly, to refuse to implement a new test, and its characteristics.
Managers also needing to know whether to support the use of a new test and, crucially, whether an old test can be safely withdrawn.
Commissioners needing to understand what to commission from laboratory services.
Anyone undertaking audit and needing standards against which to compare diagnostic practice.
Researchers needing to identify gaps in our knowledge of the utility of tests and to justify funding.
Industry needing a clearly defined route to gaining acceptance of the validity of a new test in the NHS.
IT engineers needing to develop systems, not only for 'order comms', but also expert decision support systems that encourage the appropriate and efficient use of laboratory tests.
Patients – especially people considering the use of 'direct to consumer' testing

Of course, identifying what should be done is easier than identifying how to achieve it, who should do it and how to pay for it. A collaborative approach will be needed, involving the professions, industry and Government. The volume of work required just to review the existing evidence in relation to currently accepted tests is huge; it was agreed that a system of prioritisation based on the most common care pathways would be sensible. Two possible starting points for the database were identified: the list of diagnostic tests already being developed within the NHS within 'Connecting for Health' or an expansion of *Lab Tests Online*[®], available at www.labtestsonline.org/, which at present is directed more towards patients than professionals.

An expansion of *Lab Tests Online*[®] would have the added advantage of starting to address a related

problem, not considered at this meeting. If you are a member of the public, reading an advertisement for a marvellous new test available from your pharmacist or over the internet, which will guide you to a healthy lifestyle or even tell you when and of what you will eventually die - where can you turn to obtain an impartial and professional evaluation of that test? Market forces only work if the purchaser can evaluate the quality of what's being sold. It's a separate but important question, one recently taken up by the charity Sense About Science, which has produced a 'debunking' booklet aimed at the general public, called *Making Sense of Testing*.²

**Professor Peter Furness
Vice-President and President-elect**

References

1. Furness P, Zimmern R, Wright C and Adams M. *The evaluation of diagnostic laboratory tests and complex biomarkers. Summary of a Diagnostic Summit*, 14–15 January 2008. London: PHG Foundation and The Royal College of Pathologists, 2008.
2. Sense about Science. *Making Sense of Testing: A guide to why scans and health tests for well people aren't always a good idea*. London: Sense about Science, 2008. www.senseaboutscience.org.uk (accessed 14 May 2008).