difficulty is that a procedure's clinical impact is often highly dependent on the skill of the clinician doing it. Clinicians working in settings less favourable than those from which major series have been reported will have to climb the learning curve and may not even then achieve such good results. Also, procedures may evolve and change appreciably, especially when they are new.

There may also be cultural and educational reasons why interventional clinicians do not find rigorous evaluation attractive. These may include the clinicians' decisive approach to clinical practice and their personal involvement in procedures.⁴ They may then be less willing to recognise the existence of uncertainty and undertake research to reduce it. The problem is not confined to new interventional procedures: more widely, treatments in general surgery are half as likely as treatments in general medicine to be based on randomised controlled trials.^{5 6}

The benefits and harms of new interventional procedures need to be considered together, and compared with existing treatments and no treatment. For example, laser in situ keratomileusis (LASIK) is used to treat myopia, a usually minor impairment corrected with great safety and effectiveness with spectacles or contact lenses. Even a low rate of complications might lead some patients to have qualms about the procedure. By contrast, people with an otherwise untreatable and rapidly fatal cancer might accept a procedure with a high immediate mortality after which survivors are likely to be cured.

Many innovations bring positive advantages which patients, clinicians, and healthcare systems value. However, lack of information on their benefits and harms brings unnecessary risk and may militate against their widespread use. What can be done? Firstly, clinicians need to do better research, from their first use of a new procedure. They need more training in how to evaluate their innovations, more collaboration with researchers skilled in evaluating health technologies, and more funding. Secondly, manufacturers of medical devices need to do more than the minimum required by their regulators and to show by research if their products change patients' outcomes. Thirdly, healthcare systems need to support responsible clinical innovation by providing guidance, not least to ensure that patients are fully informed before deciding whether to accept an interventional procedure.

In the United Kingdom, the National Institute for Clinical Excellence (NICE) has an interventional procedures programme (www.nice.org.uk/ip), which aims to improve care for patients by specifying the conditions under which a new interventional procedure should be used.⁷ For procedures with uncertain safety and efficacy, these include actively managing the uncertainty in a way that minimises risk. This means enhancing consent arrangements for patients so that they understand what is known and what is not; working with hospital clinical governance staff to manage the risks that the procedures may create; and monitoring outcomes by clinical audit or research. NICE guidance describes the limitations in the existing evidence, and, for selected procedures, NICE commissions systematic reviews and primary research.

No clinical guidance can remedy the underlying reasons for the limited attention given to evaluating procedures. Their potential for helping and harming patients is similar to that of pharmaceuticals, yet investment in gauging their benefits and harms is much less. Until this is addressed patients will continue to be exposed to poorly understood risks, and uncertainty will continue to hamper the safe introduction of interventional procedures into routine clinical practice.

Tom Dent *director, interventional procedures programme*
(tom.dent@nice.nhs.uk)

Sally Wortley *analyst, interventional procedures programme*

Bruce Campbell *chairman, interventional procedures advisory committee*

National Institute for Clinical Excellence, London WC1V 6NA

Competing interests: TD and SW are employed by NICE. BC's employer is reimbursed for time he spends on NICE's work.

- 1 Strasberg SM, Ludbrook PA. Who oversees innovative practice? Is there a structure that meets the monitoring needs of new techniques? *J Am Coll Surg* 2003;196:938-4.
- 2 Faulkner A, Kennedy LG, Baxter K, Donovan J, Wilkinson M, Bevan G. Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model. *Health Technol Assess* 1998;2:6.
- 3 Medical Services Advisory Committee. *Radiofrequency ablation of liver tumours: assessment report*. Canberra: Department of Health and Ageing, 2003. www.health.gov.au/msac/pdfs/reports/msac1052.pdf (accessed 7 Jun 2004).
- 4 McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ* 2002;324:1448-51.
- 5 Ellis J, Mulligan I, Rowe J, Sackett DL. Inpatient general medicine is evidence based. *Lancet* 1995;364:407-10.
- 6 Howes N, Chagla L, Thorpe M, McCulloch P. Surgical practice is evidence based. *Br J Surg* 1997;84:1220-3.
- 7 Campbell B, Maddern G. Safety and efficacy of interventional procedures. *BMJ* 2003;326:347-8.

Pharmacogenetics—expectations and reality

Drug response and toxicity depend on genes, environment, and behaviour

Thirty years have passed since Mike Rawlins, the current chairman of the National Institute for Clinical Excellence (NICE), coauthored a small but perfectly formed book entitled *Variability in Human Drug Response*.¹ In the interim we have witnessed the waxing and waning of clinical pharmacology and the inexorable rise of genomic medicine. Genomic medicine has generated many expectations with regard

to the advent of "personalised medicine" and "individualised prescriptions," fuelled by the pace of technological advances in genotyping; enthusiasts extrapolating beyond small proof of principle and retrospective studies; and a few apparent success stories, such as the treatment of breast cancer with trastuzumab (Herceptin) and of HIV with abacavir (Ziagen).

BMJ 2004;329:4-6

Examples of polymorphic enzymes and receptors, affected drugs, and unwanted responses that might be avoided or reduced by genetic testing

Enzyme or receptor	Drugs	Unwanted response
TPMT	6-mercaptopurine, azathioprine	Myelotoxicity, treatment-related secondary tumours
CYP2C9	Warfarin	Haemorrhage
CYP2C19	Proton pump inhibitors	Incomplete ulcer cure rate
CYP2D6	Antipsychotic drugs Tricyclic antidepressants	Extrapyramidal effects Confusion, cardiotoxicity
UDPGT 1A1	Irinotecan	Diarrhoea, myelosuppression
Beta ₂ -receptor	Beta ₂ -agonists	Increased airway reactivity
Ryanodine receptor	Inhalation anaesthetics	Malignant hyperthermia
Dopamine D3 receptor	Typical antipsychotic drugs	Tardive dyskinesia

Abbreviations: TPMT=thiopurine methyl transferase, CYP=cytochrome P450, UDPGT=uridine diphosphate glucuronosyl transferase.

But the promise of pharmacogenetics has largely remained unfulfilled. In general, drug response and toxicity are likely to be a complex function of the influence of many genes interacting with environmental and behavioural factors. Trastuzumab is effective in only the 15-20% of breast cancer patients who respond positively to a test for mutations in the tumour that over-expresses human epidermal growth factor receptor (HER)-2.² And about half of white, male, HIV positive patients with specific variations in the HLA-B gene are likely to develop severe reactions to abacavir.³ These examples remain controversial with regard to the specificity, exclusivity, cost, and reliability of the associated genetic testing; they also represent cases in which gene penetrance and frequency are relatively high.

Outside clinical pharmacology, poor prescribing skills, interactions between drugs and between drugs and herbs, and lack of adherence to treatment are insufficiently acknowledged as causes of therapeutic failure or adverse drug reactions. Although genetic testing is often held up as a way of improving compliance, this phenomenon has a notable behavioural component that is independent of drug and disease.⁴

Given that genetic factors need to be put into perspective, the challenge now is to assemble large, prospective, multidisciplinary, multicentre projects to assess the real clinical and economic value of predictive genetic testing in drug therapy. This coincides with the new dawn of clinical investigation ushered in by the perceived failure of new drug development and a recent flurry of position papers.⁵⁻⁶ Whereas the pharmaceutical industry seems largely to be taking a "wait and see" attitude with regard to targeted treatment solutions as opposed to the traditional "one size fits all" approach, the UK Department of Health has taken the initiative in calling for research proposals for the development of genetic tests for existing drug therapies that have a greater than 50% chance of reaching the bedside in five years' time. This may be a tall order, but it is one that focuses attention on the key requisites for evaluating the potential cost effectiveness of pharmacogenetic strategies to help healthcare providers.

Several primary characteristics that will enhance the cost effectiveness of pharmacogenetic testing have been proposed.⁷ These include severe clinical or economic consequences that can be avoided through the use of a test, difficulty in monitoring of drug

response using current methods, lack of an alternative drug with equivalent therapeutic profile and price, the existence of a well established association between genotype and clinical phenotype, the availability of a rapid and inexpensive genetic test, and a relatively high frequency of the variant gene. The table lists some examples of relatively low hanging fruit with regard to prospective evaluation, albeit with varying degrees of concordance with these criteria. A recent example of the development of a predictive dosage algorithm incorporating genetic testing showed that 39% of the variance in the maintenance dose of warfarin could be explained by a combination of genetic, clinical, and demographic factors.⁸ The use of the algorithm also more than halved the risk of adverse drug reactions. Further refinement of the model could lead to a cost effective improvement in the use of warfarin.

In the treatment of complex diseases, such as cancer and hypertension, and in the prediction of adverse drug reactions in general, the science has not advanced much beyond the fishing expedition or stone turning approach to identify important combinations of genetic determinants of drug response. Accordingly, in these areas—with the exception of the use of genetic tumour markers—the promise of relatively straightforward predictive tests is likely to be much further down the line.

Several criteria can be proposed for the conduct of prospective studies to develop predictive genetic tests of drug response. The major candidate genes should be well documented as being functionally relevant and should cover all aspects of the pharmacology and toxicology (receptors, enzymes, transporters, immunomodulators). The possibility of significant global and regional variation in the relevant genes should be addressed. Demographic and environmental factors (for example, drug interactions) should also be considered as covariates. Drug exposure and compliance should be assessed rigorously to reduce "noise" in the data. The study should have adequate power—assuming that the statisticians can agree on methods for assessing multiple gene effects. Once a predictive model with data from an initial set of patients is developed, this should be tested prospectively in a further set. A professional cost benefit analysis should be included.

Potential constraints on implementing the promise of precise prescriptions include general ethical and racial issues,⁹ the disincentives to clinical investigation imposed by recent inflation in the associated paperwork, costs involved in gaining ethical and regulatory approval, and considerations of patients' and doctors' acceptance of genetic testing.¹⁰ Additional cost implications may occur in countries where healthcare is not free at the point of delivery. Historically, methods of predicting drug responses, such as renin profiling and acetylator phenotyping, have had a very low uptake in the United Kingdom. Genotyping may be different? Time, as ever, will tell.

Geoff Tucker *professor of clinical pharmacology*

Academic Unit of Clinical Pharmacology, Pharmacokinetics and Pharmacogenetics Group, University of Sheffield, Royal Hallamshire Hospital, Sheffield S10 2JF (g.t.tucker@sheffield.ac.uk)

Competing interests: None declared.

- Smith SE, Rawlins MD. *Variability in human drug response*. London: Butterworths, 1973.
- Vogel CL, Cobleigh MA, Tripathy D, Debu G, Harris JC, Lyndsay N, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002;20:719-26.
- Hetherington S, Hughes AR, Mosteller M, Shortino D, Baker KL, Spreen W, et al. Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. *Lancet* 2002;359:1121-2.
- Urquhart J. The odds of the three nons when an aptly prescribed medicine isn't working: non-compliance, non-absorption, non-response. *Br J Clin Pharmacol* 2002;54:212-20.
- Food and Drug Administration. *Innovation or stagnation? Challenge and opportunity on the critical path to new medical products*. Washington: FDA, 2004. <http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.pdf> (accessed 15 Jun 2004).
- Academy of Medical Sciences. *Strengthening clinical research*. London: Academy of Medical Sciences, 2003. www.acmedsci.ac.uk/p_scr.pdf (accessed 15 Jun 2004).
- Veenstra DL, Higashi MK, Phillips KA. Assessing the cost-effectiveness of pharmacogenomics. *AAPS PharmSci* 2000;2:E29.
- Gage BF, Eby C, Milligan PE, Banet GA, Duncan JR, McLeod HL. Use of pharmacogenetics and clinical factors to predict the maintenance dose of warfarin. *Thromb Haemost* 2004;91:87-94.
- Lee SS-J. Race, distributive justice and the promise of pharmacogenomics: ethical considerations. *Am J Pharmacogenomics* 2003;3:385-92.
- Hapgood R. The potential and limitations of personalised medicine in primary care. *Br J Gen Pract* 2003;53:915-6.

Collecting and sharing information about harms

Appropriate strategies to communicate this information are essential

Prescribing a drug represents a trade-off between its benefits and harms. The harm to benefit ratio varies according to the condition being treated, the drug's pharmacology, and the availability and safety of other therapeutic options. A decision to prescribe a drug therefore has to be made on the basis of all the information available. There is, however, often an imbalance in such information. Although high quality data regarding benefits usually come from randomised controlled trials, systematic reviews, and evidence synopses, information on harms may be limited. This is particularly the case for new drugs, which usually have been given to only a few thousand patients, and often for relatively short periods before licensing.¹ This allows common adverse events to be detected, but not rare events or those that occur only with prolonged exposure. Consequently, established compounds continue to cause much harm²—for example, approximately 5% of all admissions to US and European hospitals are due to adverse drug reactions.^{w1} A much publicised meta-analysis showed that such reactions caused more than 100 000 deaths in the United States in 1994.^{w2}

Monitoring harms after launch of a product is thus crucial, since regulatory action may be necessary to protect public health. Systems to monitor drug safety are well developed in the United States and most EU countries (but hardly at all in developing countries). For example, the probability that a new drug will acquire a "black box warning" or be withdrawn in the United States has been estimated to be 20% over 25 years.³ In the United Kingdom 4% of all licensed products are withdrawn after marketing because of safety problems.⁴ However, a great number of patients have to be treated before an uncommon adverse event is detected and regulatory action taken. For example, 7.5 million patients were exposed to terfenadine before regulatory action was instituted in relation to its QT prolonging effects.⁵

Postmarketing surveillance of drugs uses many methods.^{w3} Spontaneous reporting systems—such as the yellow card scheme in the United Kingdom, which is 40 years old—have many strengths, as highlighted by the recently published yellow card review.⁶ Spontaneous reporting schemes should be regarded as hypothesis generating tools as any signals need to be confirmed by pharmacoepidemiological studies.^{w4} A great limitation of all reporting systems is under-

reporting: only 10% of serious adverse drug reactions are reported.¹ This has led to a widening of the reporter base—for example, by including pharmacists⁷ and nurses.⁸ Much inconsistency exists, however, in reporting systems in the European Union and worldwide,^{w5} which need to be harmonised. The yellow card review has recommended the introduction of direct reporting by patients in the United Kingdom.⁶ Although reporting by patients may be valuable when intensive monitoring is taking place or when reporting is through an intermediate healthcare professional,^{w6} it is unclear whether this is also true of direct patient reporting, which needs robust evaluation. The aim of any spontaneous reporting system should be the prompt detection of new drug safety issues. The expansion of the reporter base may contribute, but it is crucial that this does not occur at the expense of an increase in the ratio of signal to noise. Perhaps improved data mining techniques that use statistical techniques such as cluster analysis, link analysis, deviation detection, and disproportionality assessment will be able to cope with "noise" and enable earlier detection of safety signals in databases of information on health care.⁹

There are several reasons to be optimistic that data on harms will improve over the next few years. Firstly, inadequate reporting of adverse event data in randomised controlled trials and the difficulty in identifying information on harms in literature databases are now widely recognised,^{w7} and are likely to lead to specific reporting requirements in the CONSORT criteria. Secondly, systematic reviews have had a major influence on the assessment of benefits of medicines, but not of harms, not least because most data on harms are collected from observational studies and case reports, not from randomised controlled trials. The intention of the *BMJ* and the Cochrane Collaboration to improve the reporting of harms in their publications is certainly a step forward.¹⁰ Thirdly, the International Conference on Harmonisation has published guidelines for pharmacovigilance planning (E2E),^{w8} which will allow more structured assessment of harms after licensing, and include investigation of drugs in groups that may be at particular risk—for example, elderly

Additional references w1-w11 are on bmj.com

BMJ 2004;329:6-7