

The limitations of evidence

Karl Popper's most distinctive contribution to the epistemology of science was his contention (itself an hypothesis) that sound scientific knowledge is based on the development of testable hypotheses, that may then be falsified by the new evidence that emerges from the process of testing. His *Logik der Forschung* (1934) has been both translated¹ and elucidated for the general reader² by Popper himself, and vested with elegance and plausibility by Peter Medawar in *Pluto's Republic*³. Imaginative hypothesis followed by meticulous testing for concordance with what is already known, or ascertainable by trial (Popper's 'hypothetico-deductive process') is probably closer to the reality of scientific discovery than the alternative model of Baconian induction from raw facts. It is also appealing, in asserting the importance of imagination in science, in contrast to the 'dry-as-dust' stereotype that is all too prevalent among those whose acquaintance with actual scientists is limited. The story of science blazes with what Michael Polanyi called 'passion and controversy'⁴. Moreover, Popper's model emphasises the provisional or contingent nature of scientific knowledge – even a Newton awaits his Einstein, and any beautiful theory is vulnerable to an ugly fact. A scientist who embraces *hubris* is living beyond his intellectual means.

Nevertheless, total surrender to the scepticism that is theoretically demanded of the scientist is in itself unwise; there are degrees of probability so high that pragmatically, though not in absolute theory, they carry virtual certainty – as Henry Harris puts it: 'I do not believe that it will ever be shown that the blood of animals does not circulate; that anthrax is not caused by a bacterium; that proteins are not chains of amino-acids'⁵. Of course, these examples are matters of direct observation (once you know how), and not to be confused with the broad paradigms that concern the philosopher of science, and which are subject to radical changes, or 'revolutions'⁶ at (fortunately) long intervals. Active scientists are in general indifferent to the origin and even the validity of the conceptual framework that they accept and within which they work – as Medawar pithily puts it: 'the history of science bores most scientists stiff. A great many highly creative scientists... take it quite for granted, though they are usually too polite or too ashamed to say so, that an interest in the history of science is a sign of failing or of unawakened powers'³. Happily, a flawed framework does not prevent all accretion of knowledge through observation. The Greeks could add to physical science within a framework torn between Leucippus' atoms in a void and the four elements of Empedocles; somewhat

sadly, the invention of gunpowder did not have to await either the theory of phlogiston or the discovery of oxygen.

Those who seek to advance medical science, or to practise competent medicine, are doubtless quite as likely to be detached from the evolution and validity of their basic assumptions as those in other branches of science or professional activity. So far as effectiveness of medical practice is concerned, this may not in earlier times have greatly mattered; over many centuries, the faithful followers of Galen adhered to a framework of the four humours, without any specific baleful effect on what they did or failed to do. But in that respect, things have changed. In many illnesses there are great benefits from giving the right treatment, and corresponding dangers from giving the wrong treatment or even from omitting treatment, which in days gone by was often the right thing to do. The sheer power and specificity of the drugs now available, the frequency of side effects, and the multiplicity of indications for their employment, demand the utmost responsibility in their use, a responsibility whose recognition has fostered the discipline of clinical pharmacology, an essential component of undergraduate and postgraduate medical education.

A more recent manifestation of this proper concern for the best possible deployment of remedies may be seen in the advocacy of 'evidence-based medicine' (EBM)⁷. In his 1996 Office of Health Economics Lecture⁸, Professor David Sackett had an opportunity to describe EBM in a less formal and more extended way than would be appropriate in a journal article. For example, he is able to say: 'I'm a student of neither ethics nor health economics, and my knowledge in these areas is mundane'. But he seems to me to have the root of the matter when he says: 'although I've drawn salaries from universities, governments, and now from the NHS, I've always considered my real employers (ie those with the highest call on my loyalty) to be my patients, my students, and the junior doctors on my clinical teams'.

Sackett defines EBM as 'the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients'. This involves 'integrating individual clinical expertise with the best available external clinical evidence from systematic research'. Individual clinical expertise derives from the care of patients and from accumulated personal knowledge of the natural history of disease – the medicine of Sydenham and Ryle. External research-based clinical evidence is derived in part from 'the basic sciences of medicine', but 'especially from patient centred clinical research into the accuracy and precision of diagnostic tests (including the clinical examination), the power of prognostic markers, and the efficacy and safety of therapeutic, rehabilitative and preventative regimens'. Although these types of relevant evidence range far beyond that derived from randomised trials and meta-

SIR DOUGLAS BLACK MD FRCP, Past President, Royal College of Physicians of London

analyses, Sackett allots a special place to that type of evidence as being 'so much more likely to inform us, and so much less likely to mislead us' than the 'non-experimental approaches', which in his view 'routinely lead to false-positive conclusions about efficacy'. Indeed, he regards evidence derived from the systematic review of several controlled trials as the 'gold standard' for judging whether a treatment does more good than harm. This is the message first proclaimed by Archie Cochrane⁹, whom to know was to love, but whose manifold gifts did not include inerrancy.

Sackett's description of what is meant by EBM is framed in terms of individual patient care; but the principle of using best available evidence is also relevant to 'group decisions' on the provision of services, in which 'the overriding objective... is that of maximising the total cost-utility in that group'. Both in individual medicine and in public health medicine, clinical judgements have to be modulated to an appropriate extent by considerations of ethics¹⁰ and of health economics¹¹.

Comment on EBM

There can surely be no difficulty in accepting the general proposition that medical practice and the provision of health care should be based on the best available evidence, or in welcoming the complementarity accorded by Sackett to 'individual clinical expertise' and 'external clinical evidence from systematic research'. Any movement which encourages self-scrutiny and self-criticism has to be a good thing, so long as it stops short of aborting that modicum of self-confidence which enables us to live and practise, while aware of the strong element of uncertainty that is inescapable in all branches of health care. However, since criticism is the acid that etches medical systems into better shape, I would raise one particular concern – the primacy given to 'RCT evidence' – and two more general ones: the validity and applicability of 'group evidence', and the extent to which the 'EBM principle' can be applied to the totality of medical practice.

Relative value of RCT evidence

The development of the randomised controlled trial (RCT) by Bradford Hill and Richard Doll is both theoretically and practically a milestone in medical history. It was a brilliant response to the increasing problems set by the proliferation of agents that are highly effective but also potentially hazardous from their side effects; it has in many clinical contexts provided sound information not only on which drug to use but also on how much of it to give. These are immense gains, but to the extent that Cochrane and Sackett imply a primacy over other sources of therapeutic evidence, I think they take a step too far, and one that would perhaps take them beyond Doll and Hill themselves.

The quality of evidence should be assessed not by the method by which it is obtained but by its strength or weakness. RCTs can produce evidence of great strength (and

have often done so) when a clinical situation is: common, so that the trial can be carried out in one centre; easily defined so that inclusion and exclusion are simple; and has little variation between patients. Whilst there are ways of organising multi-centre trials that largely preserve their validity (at some cost in increased complexity), the most important derogation to the strength of RCT evidence stems from patient variation and imperfect clinical taxonomy. When there are many variables with an effect on outcome, as in coronary thrombosis, or when an apparently simple label conceals taxonomic complexity, as in the nephrotic syndrome¹², then the results of RCTs tend to be inconclusive or conflicting, though probably still stronger as evidence than the results of unsystematised clinical observation of the effects of therapy.

There are other risks in the 'systematic review of several controlled trials', commended 'especially' by Sackett⁸ in comparison with the single trial. It is true that the influence of one trial, idiosyncratic for whatever reason, may be 'diluted out' by numerous others. But as against that, occult variations in procedure, such as may lie hidden even in a single multi-centre trial, must be almost inevitable in a group of trials subjected to meta-analysis. While trials can be scrutinised for gross errors of methodology, and excluded if such are found, less obvious but still material differences in procedure may not be apparent to a scrutiny diffused over a number of protocols. Another problem, particular to the aggregation of trials, is the systematic bias that arises from the reluctance of authors to report, and of editors to accept for publication, trials with a 'negative' result, which may be just as relevant as a 'positive' trial in forming a true picture¹³.

In summary, a well-conducted RCT on a clearly defined issue can yield evidence of a probability approaching certainty, and certainly greater than that derived from unstructured clinical scrutiny of the same issue. However, there remain issues of clinical importance whose complexity makes them, for the present, 'insoluble' by the RCT route, and whose solution must await further resolution by conventional clinical or pathological analysis. There is not here a hierarchy of methodological esteem, merely an order of difficulty, and the RCT must not take 'gold standard' precedence over all other methods of clinical investigation.

Validity and the applicability of 'group evidence'

In this section I use 'group evidence' as a shorter equivalent of Sackett's 'best available external clinical evidence from systematic research'. Evidence relevant to clinical decision-making is derived from many sources: the basic medical sciences; the social sciences; the science of the 'seats and causes' of disease; the observed natural history of disease; epidemiology; the skills of diagnosis, of which speaking with the patient is paramount; and the range of possible therapies, graded for practicability and efficacy. From such a variety of sources, it is likely that the evidence will be variable in validity and relevance, as well as being potentially

overwhelming by its very extent. It is all the more important that the search should at least start in the right place, by that characterisation of the actual problem that we call diagnosis. The old-fashioned danger of being sent to the wrong specialist has its informatic equivalent of being entered in the wrong algorithm, and undergoing a series of tests appropriate to someone else.

Accurate definition of the actual problem does not remove the entire risk in applying the evidence derived from group studies to the problems of an individual. Leaving aside possible flaws in the evidence itself, its applicability depends on the individual being conformable to the group in all relevant aspects. And 'relevant aspects' may be legion, for example in the risk factors relevant to outcome in coronary arterial disease.

Pragmatic limits of EBM

Even to think of disparaging an evidential approach must be the cardinal sin of an academic physician. My backsliding had an unlikely venue in the course of a meeting convened to give advice on the permanent vegetative state; the definition and management of this are not free from the form of uncertainty to which the term 'grey area' is commonly applied. During the discussion of these matters, we were told that 'all advice coming from the Department of Health must be evidence-based'. That was certainly a boldly-stated ideal but in that context, and in many others, it scarcely corresponds to reality, unless one stretches 'evidence' to include the whole area in which we are guided by that elusive quality, common sense. Wrestling with the mechanisms, other than strict logic, by which the mind can attain certitude, Newman invoked an 'Illative Sense'¹⁴, which appears to be a faculty of correct reasoning from facts presented to us in apparently random fashion. Of course, in practical medicine our concern is rarely with certitude, but almost always with a degree of probability sufficient to justify the action that we propose to recommend; but in recognising that degree of probability, I suggest that we are more commonly persuaded by a balance of likelihoods than we are driven forward by the iron laws of evidence.

The relative frequency of 'evidence-based' and 'more-broadly-based' decisions is a matter of speculation, but it is important to recognise that many of the situations with which medicine is concerned lie outside the realm of 'scientific medicine'. Even within clinical medicine, the majority of acute situations are self-limited, irrespective of treatment, and that has been an impartial bulwark both to historical and to complementary medicine. At the other extreme, there are illnesses that inevitably lead to death or incapacity. But there are also, happily, a minority of situations in which scientific medicine makes all the difference between a cure and disaster, and this minority grows steadily larger. In our legitimate concentration on those areas in which the treatment given is critical, we must keep in mind that they are still, however important, only a small

part of the whole province of medicine, and that we have a 'duty of care' as well as a 'duty of cure'. EBM can contribute to the discharge of both of these, but it is far from being the only player on the field. Although I am no great lover of the taxonomy that distinguishes the 'science' from the 'art' of medicine (believing them to be interwoven activities), the point that I am trying to make has been better put in those terms by Robert Platt, a great clinician who taught me a great deal: 'However far the science of medicine and surgery advances, the art of medicine will remain: the art of first identifying the patient's problem (which is something more than merely diagnosing his disease) and the art of applying the science to the needs of the individual patient'¹⁵. If that is a sound critique for the practice of clinical medicine, how much more does it apply to the more general application of medical knowledge and skills which we designate as 'public health medicine'? Perhaps it is there above all that the fruits of existing evidence must first be critically examined, and then supplemented by insights derived from imaginative reasoning.

Conclusion

My intention in this paper is not to disparage evidence-based medicine, but to deprecate any attempt to equate it with the whole of medicine. I recognise that no such attempt is being made by responsible advocates⁸, but a prophet is not in control of disciples who might deny the value of medical thought and experience in situations in which a categorical evidence base is still lacking. To deny that we must use evidence to the farthest extent that it can take us would be to be guilty of obscurantism and of 'anti-science' – charges which I may perhaps be fortunate to escape, if indeed I do. What I seek to emphasise is the importance of the prior intellectual analysis of the problem, be it clinical or organisational, in such a way as to define the type of evidence that is going to be relevant. An unstructured search for evidence may only lead to confusion. Of course, intellectual analysis is fallible, and the search for relevant evidence may have to change direction, but at all stages the search has to be planned, and not random.

When all relevant evidence has been gathered, and with its help the problem is 'understood', the transition from knowledge to action may indeed be assisted by evidence derived from comparable situations, and such must be sought. However, comparability is not identity, and treatment or planning need continued intellectual review, including the planned search for retrospective evidence, without which plausible errors can lead to preventable disasters for individuals or groups.

References

- 1 Popper KR. *The logic of scientific discovery*. London: Routledge & Kegan Paul, 1959.
- 2 Popper KR. *Conjectures and refutations*. London: Routledge & Kegan Paul, 1963.

- 3 Medawar PB. *Pluto's Republic*. Oxford: Oxford University Press, 1984.
- 4 Polanyi M. Passion and controversy in science. *Lancet* 1956;**270**: 921-5.
- 5 Harris H. Rationality in science. In: Heath AF (ed). *Scientific evolution*. Oxford: Oxford University Press, 1981.
- 6 Kuhn TS. *The structure of scientific revolutions* (2nd edn). Chicago: University of Chicago Press, 1970.
- 7 Sackett DL, Rosenberg WNC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *Br Med J* 1996;**312**:71-2.
- 8 Sackett DL. *The doctor's (ethical and economic) dilemma*. OHE annual lecture. London: Office of Health Economics, 1996.
- 9 Cochrane AL. *Effectiveness and efficiency: random reflections on health services*. London: Nuffield Provincial Hospitals Trust, 1972.
- 10 Gillon R, Lloyd A. *Principles of health care ethics*. Chichester: John Wiley & Sons, 1994.
- 11 Williams A. *Being reasonable about the economics of health*. Cheltenham: Edward Elgar Publishing, 1997.
- 12 Black DAK, Rose GA, Brewer DB. Controlled trial of prednisone in adults with the nephrotic syndrome. *Br Med J* 1970;**3**:421-6.
- 13 Grahame-Smith D. Evidence based medicine: Socratic dissent. *Br Med J* 1995;**310**:1126-7.
- 14 Newman JH. *An essay in aid of a grammar of assent*. London: Longmans Green & Co, 1870.
- 15 Platt R. *Private and controversial*. London: Cassell & Co, 1972.

Address for correspondence: Sir Douglas Black, The Old Forge, Duchess Close, Whitchurch-on-Thames, Reading RG8 7EN.

ROYAL COLLEGE OF PHYSICIANS OF EDINBURGH

The Sir James Mackenzie Centenary

Atrial Fibrillation in Hospital and General Practice Consensus Conference

3 & 4 September 1998

to be held at New Hall, University of St Andrews, St Andrews

A Consensus Panel, chaired by Dr Kim Fox, Consultant Cardiologist, Royal Brompton Hospital, London, will address the following questions during a two-day conference, with the purpose of reaching a consensus.

- how do we maintain sinus rhythm in paroxysmal atrial fibrillation?
- how do we restore sinus rhythm in persistent atrial fibrillation?
- how do we achieve optimal cardiovascular function in atrial fibrillation?
- how do we prevent thromboembolism in atrial fibrillation?

Speakers to include:

Dr Ale Algra; Dr Peter Bloomfield; Professor John Camm; Professor Ronnie Campbell; Professor Harry Crijns; Dr Wyn Davies; Dr Francis G Dunn; Dr Clifford Garratt; Dr Barclay Goudie; Professor Robert Hart; Professor John Hampton; Professor Richard Hobbs; Dr Peter Humphrey; Dr Kevin Jennings; Dr Richard Lindley; Dr Gregory Lip; Dr Graham Turpie.

Abstracts are invited and should be submitted by 31 March 1998

Registration and Abstract forms can be obtained from:

Mrs Margaret A Farquhar, Consensus Conference Co-ordinator,
Royal College of Physicians of Edinburgh, 9 Queen Street, Edinburgh, EH2 1JQ
Tel: +44 (0) 131 225 7324 Fax: +44 (0) 131 220 4393 e-mail: m.farquhar@rcpe.ac.uk