CONTROVERSY

Some reservations about clinical guidelines

Anthony Hopkins

Clinical guidelines are the current fashion. There are perhaps four main grounds for this. First, the realisation that the results of clinical research are not taken up promptly and translated into everyday clinical practice. For example, meta-analysis of all the trials for thrombolysis for myocardial infarction showed that there was good evidence that thrombolysis reduced mortality by the mid-1970s, although thrombolysis only came into everyday use about 15 years later.1 2 A further example is the continued employment by some surgeons of mastectomy when it has been shown that this is not superior, in terms of survival, to local excision of the tumour and radiotherapy, and some women do not appear to have the opportunity of choosing an operation which conserves the breast.3

Next, there is an enormous variation in practice for the management of common clinical disorders. This was first reported in the mid-1930s in the UK in relation to tonsillectomy.4 Wennberg et al and Leape et al have more recently drawn attention to the wide variations in practice in the USA in the management of, for example, prostatism and coronary artery disease, upper gastrointestinal endoscopy and carotid endarterectomy.5 6 A further example nearer home in the UK is the wide variation in referral rates to hospital consultants from general practices serving similar populations.7 Some users of health services are better informed than health professionals about the extent of variations in clinical practice. Women in particular became concerned in the 1970s not only about variations in birthing practice and in the management of their common cancers, but also about the lack of opportunity to discuss various options for managing the same condition.

The third origin of clinical guidelines lies in attempts at control of the costs of health care. Variations in practice, such as length of stay after herniorrhaphy, remain considerable. Most herniorrhaphies are now performed as day cases, but other apparently similar patients are still admitted for several days. One of the major UK health insurers (the British United Provident Association) has started informing individual surgeons of how they stand in relation to the average and median lengths of stay for various procedures. Although this emphasis is on cost control, Mechanic has emphasised that one goal of guidelines is to ensure that patients receive the range and variety of services they need, and from which they can choose.8

The fourth origin for clinical guidelines relates to clinical audit. The Royal College of Physicians set up a working party on medical audit in 1988,9 and its report was published within a few weeks of the government’s white paper on the reform of the health service, Working for Patients.10 An integral feature of the white paper was the requirement to develop medical audit. Indeed, Kenneth Clarke, the then Secretary of State for Health stated that ‘medical audit, where doctors analyse and improve performance through peer review, would have more impact than any other element of the health service reforms’.11 It rapidly became clear that audit could only take place against some sort of standard. Clinicians needed to develop standards in consultation with those actively concerned in research into the process or outcomes of care that was to be audited. Paediatricians have helped the Research Unit of the Royal College of Physicians prepare guidelines on the management of urinary infections in childhood,12 the management of convulsions with fever,13 the management of the acute nephrotic syndrome in childhood,14 and the management of the respiratory distress syndrome.15 The British Paediatric Association has since developed its own research unit, under the direction of Professor David Baum. As time has passed, the word ‘standard’ has been dropped, largely on account of anxieties, misplaced in my view, that the use of the word standard implies a greater liability to litigation for negligence if a physician’s care deviated from that standard. In practice, the position is as it has always been – that if a doctor follows a course of action that is supported by a responsible body of his peers, then it is most unlikely that the courts would consider any action to be negligent, even if it did not follow a standard. The legal significance of guidelines has recently been fully considered by Hurwitz16 and the Clinical Outcomes Group (the Department of Health committee which consider matters related to audit) has recently set up a subgroup to consider further the legal issues.

As the last few years have gone by, more and more groups are writing guidelines, and yet I am bound to say that in my experience, clinicians are far more eager to write guidelines than to pilot audit protocols developed from those guidelines. To my mind anyone who participates in the preparation of a guideline must be prepared also to help develop an audit protocol based upon that guidelines, and,
moreover, be prepared to pilot that audit protocol for ease of understanding, for reliability, and for validity in his or her own unit before publishing such a protocol.

It is of course easy enough to write a guideline of some sort by pulling together from textbooks of medicine, scientific reviews, and peer judgment a body of opinion that is more or less sensible. However, the standard of work on guidelines, as in other branches of medicine, is continually rising. The work of the Oxford Perinatal Epidemiology Unit and the subsequent development of the Cochrane Centres has shown how rigorous structured reviews of the scientific medical literature need to be in order to be reliable and valid. For example, as there is a bias towards publication of positive results of trials rather than negative results, there is a tendency for an intervention to appear to be more effective than it really is, if only the published trials have been reviewed. Furthermore, randomised controlled trials of an intervention are usually based on a carefully selected and defined population, often with an upper age limit, the intervention being carried out in university hospitals with research staff to increase patient compliance, and with senior medical and surgical staff with a high degree of technical skill. The results cannot necessarily be extrapolated to the general population with the disorder, looked after by the ‘average’ physician or surgeon and ward staff. The population subsequently treated may well be of a broader age band than in the trial and often with co-morbidities. We should therefore do our best to study the effectiveness of the intervention in the general population before recommending the wide adoption of a new treatment.

Parallel with this increasing rigour of review is increasing expense. The early guidelines prepared by the Research Unit of the Royal College of Physicians in association with specialist groups such as the British Paediatric Association cost about £3500 – travelling expenses, a good lunch, and the costs of publication. It is not unusual for the development of guidelines published by the Agency for Health Care Policy and Research in the US to cost in excess of $200 000. The guidelines that are cheap to prepare tend to be judged by the same standards as those which cost a great deal, which is painful, but probably wise. The NHS Executive has recently commissioned the development of criteria by which the rigour of guidelines developed in the UK can be appraised.

Without suggesting that we should follow the US example of heavy expenditure, I have some views about guidelines which may be helpful to share with those who wish to consider writing them. I have been guilty of most of the following solecisms and can claim no special expertise in this area, but I hope that each guideline that the college develops, in association with specialist societies, is slightly better than the last. In this paper, I am not addressing the considerable difficulties in changing practice through the implementation of guidelines, ably covered elsewhere.

(1) There is an understandable tendency to write guidelines for a clinical diagnosis, rather than for a presenting problem. For example, guidelines were written for the management of angina or myocardial infarction, rather than for the management of chest pain. Guidelines were written for the management of children with febrile convulsions, rather than for the management of a convulsing child with fever, who might conceivably have meningitis. It is undoubtedly much easier to write guidelines for clearly defined diagnostic categories, but the product is less likely to be of help to young doctors in training, who are faced with clinical problems before diagnosis. Guidelines written for a ‘diagnosis’ are also less likely to be useful for audit. It should, in theory, be easy enough to retrieve the notes of all the patients discharged with a diagnosis of say, myocardial infarction, as these should be picked up by the hospital coding system. But what of the patient who had a dissecting aneurysm that was mismanaged as a myocardial infarction for the first 12 hours of his care, or the patient with a functional chest pain who was overinvestigated with coronary angiography? In short, I believe the guidelines should address clinical problems, not clinical diagnoses. Good examples are those admitting diagnoses used by physicians responsible for the care of elderly people – ‘found lying on floor’ or ‘recently gone off her legs’. Patients and their relatives also have no difficulty in deciding what a ‘problem’ is. I believe that most medical readers will recognise the truth behind these sentences, but they do of course run contrary to the hope of the NHS Executive and of health insurers that if only doctors treated patients with standard diseases in a standard way, then the results would be better and costs less.

(2) The target audience of health professionals in association with guideline writers is often not adequately identified in many published guidelines. By the nature of academic medicine, doctors working in university medical and hospital practice are more abreast of current advances in research and in effectiveness than those working in primary care. Guidelines written for the appropriate management of patients admitted acutely to hospital may well be quite inappropriate for the management of the same patient a few hours earlier in the patient’s home. A whole raft of different possibilities for appropriate management emerge once the patient has been referred to hospital. There is some recognition of this now, with the concept of ‘shared care’, but general practitioners did not find useful many of the guideline statements for asthma written largely by specialists.

(3) The target population of patients for whom the guidelines are written is also often not clearly identified. It is hard to avoid writing dogmatic statements which cannot possibly affect all patients. For example, in guidelines relating to the management of stroke, many suggested interventions such as early physiotherapy and language therapy may be quite inappropriate if applied to an elderly lady living in a nursing home who has already had five
previous strokes, with considerable impairment already of mobility and language. Yet it is difficult to target populations without writing in exclusion causes which imply that some patients are 'worth less' than other patients. The point I am making here is that doctors are being left in the air as to which patients the guidelines should be applied.

(4) The core feature of clinical guidelines is that they should be based upon good scientific evidence that the recommended intervention does actually improve outcome. Most such evidence now comes from randomised controlled trials, although it is increasingly realised that, for trials to have sufficient power, they must often involve large numbers of subjects. However, many physicians deal with chronic disorders and the results of aging, for which the technical interventions that are available do not influence outcome very much. To take an example from my own specialty of neurology, most patients with multiple sclerosis want to see a neurologist. They want to have the diagnosis made as firmly as it can be, to obtain some idea about their future, and to discuss the various options for treatment. However, the evidence of the effectiveness of any intervention in the management of multiple sclerosis, in terms of slowing the progress of the disease or shortening the duration of a relapse, is, with the exception of short term benefit from methylprednisolone, extremely scanty. Much of the work of many specialists falls into the categories of support, reassurance, and explanation during the evolution of a chronic illness or of the aging process. To take again an example from the specialty of paediatrics, a paediatrician will contribute most to the management of a child with cerebral palsy by explanation, advice, and support over several years. The individuality of the family circumstances and of the child's particular disability are such that it is unlikely that any technical guideline would be very useful. Clinical guidelines are not suitable for the management of such patients, in which patients and their families have to modify their lives to accommodate the illness, rather than expect technical relief. In short, clinical guidelines are largely concerned with technical aspects of care, yet the general public, although primarly concerned with this, do remain, particularly in the example of chronic illness, concerned about the interpersonal aspects of care, and continuity of care.

Where there is no research based evidence of effectiveness, then useful guidance can be given that reflects current ethical and professional views. However, a clear cut distinction must be made in any publication between views that reflect professional consensus or the received ethical view from views that are based upon evidence provided by research. For example, if guidelines were written in relation to in vitro fertilisation (IVF), guidance might be written about the age of the mother to whom IVF might be offered – but this would reflect an ethical judgment and a judgment of cost effectiveness rather than published evidence of lack of efficacy.

(5) There is a potential confusion about the outcome which a guideline is targeting. For example, there are some guidelines about stroke indicating that patients with a stroke really 'ought' to have computed tomography. Although clearly effective in improving the distinction between haemorrhage and infarction, and valuable for categorisation of patients recruited to stroke trials, there is as yet no evidence that any imaging study influences stroke outcome – the final arbiter of effectiveness.

(6) Although guidelines are written by health professionals, users of health services should be the arbitrers of the outcomes for which guidelines are written. For example, between about 1965 and 1980 technical issues in obstetric care were considered to be preeminent, and hospital antenatal care and confinement were the norm. An increasing number of women decided that, although primarily continuing to be concerned about the safety of their child, the experience of childbirth under their own control was an outcome of pregnancy they valued highly. Returning also to the example of breast surgery given above, the conservation of breast tissue was valued more highly by most women than by many (predominantly male) breast surgeons, who valued more highly radical eradication of the tumour and glands. When considering evidence of effectiveness, therefore, one has to consider the effectiveness of the intervention balanced by outcomes valued by patients. It follows that the experiences of patients who have been through an illness or a procedure should be brought into the consultative process at an early stage before guidelines are drafted. Clinical guidelines must reflect the values of users of health services, recommend processes of care that are acceptable to the population, and offer sufficient flexibility to take into account the preferences of individual patients.

(7) The next point relates to the clarity of guidelines. If a guideline states that such and such an intervention is "useful", then the reader is left in doubt as to the evidence about the circumstances in which the investigation should be performed. I have myself tried to be strict about this point, but to my chagrin, I found after publication (jointly with my colleague Professor David de Bono on guidelines for the management of stable angina) that the phrase 'Further investigation might include ...'. It is remarkably difficult to avoid such conditional statements.

(8) Most guidelines are written for patients with isolated disorders – a hernia, a stroke, a myocardial infarction. However, the reality of medical practice in the average district general hospital in the UK is that nearly 90% of medical patients are admitted as emergencies, and more than two thirds are over the age of 65. By the accrual of age, many patients have more than one active illness (co-morbidity), so that management has to be tailored to fit not only the most active component of a principal problem, but also other ongoing medical and social problems. Here perhaps is a hint of a return to the necessary 'art' of medicine, when
considering how to apply guidelines written for a population with the disorder to the special circumstances of an individual patient.

(9) This last point is important also in relation to the use of guidelines for audit. If audit is to be based upon outcome (and increasingly we hear of purchasers considering placing contracts with units which achieve the most favourable outcome), then the problem of co-morbidity ('case-mix severity') must be addressed. Even using such a comparatively simple outcome as mortality, the Health Care Financing Administration in the US has shown that the more it corrects its mortality figures for co-morbidities as reflected in the simple measures provided by claims data, then less and less become the differences in mortality for various conditions between hospitals. There is plenty of research evidence to show the factors that are associated with poor outcomes, not only in malignant disease, but in cardiovascular and cerebrovascular disease, and indeed in most specialties. In coronary disease, it is known that survival is associated with left ventricular function, so it follows that if outcomes of myocardial infarction in different hospitals are to be compared, then this variable must be recorded in every case. In practice, the better and more detailed the research, the more and more variables which influence outcome can be identified. One of the challenges of outcomes research is to attempt to determine for some of the principal disorders of interest to the public those measures of case-mix severity which can be cheaply and reliably collected, and which can be used for correcting crude outcome figures.

(10) Until about one year ago, I and I believe most of my colleagues in the UK concerned with the development of clinical guidelines considered that we were working to improve medical care by making available evidence about effective clinical practices, and by providing a basis for clinical audit. Our principal target audiences for such guidelines therefore were consultants, general practitioners, and doctors in training. In December 1993, the NHS Management Executive published an executive letter to general managers and directors of public health informing them of 'A developing initiative to integrate professional guidelines more effectively into the delivery of health care'. In the executive letter, the Health Care Director and Director of Nursing of the NHS Management Executive wrote 'We now aim to work with the professions to identify and develop guidelines which will be useful in informing discussions between purchasers and providers on the development of service specifications and contract negotiations'. However, there is an enormous gap between guidelines written for health professionals, and guidelines to be used in the contracting process. It is unrealisable to consider that a guideline which goes so far as to specify the exact dose of a drug that research has shown to be effective in a certain combination of circumstances should form part of the contracting process, although it could be a useful point for clinical audit. This gap between guidelines that are suitable for health professionals and guidelines that are suitable for the contracting process has not yet been adequately recognised. If guidelines for health professionals are to be translated, as it were, into the contracting process, then it is essential that a subset of data items are identified that can be cheaply and reliably collected, and which might be used for monitoring the performance of a contract. The present results and expense of clinical audit suggests that only a limited number of data items can be collected about each episode of illness. I am concerned that purchasers may require vast amounts of information about different processes of care, this information proving expensive to collect, and on proper examination more or less valueless for comparing outcomes between clinical teams. There has, for example, been little research as yet in the UK into the reliability of clinical audit (whether or not two different observers, going back over some recorded data, can retrieve the same data items), yet there is already considerable evidence of bias through selective retrieval of records.

Although the executive letter commended purchasers a number of guidelines that have already been published, including two with which I myself have been concerned, in my view none are suitable for the contracting process. For those guidelines already developed, the whole process will have to be gone through again with this different audience in mind. However, it should be acknowledged that guidelines, even as they stand, may help purchasers resist less well informed local personal opinion and local pressure groups.

(11) In the last financial year, the NHS has fused the budget allocated for medical audit with the budget available for the audit of the care given by other health professionals such as nurses, clinical psychologists, physiotherapists and so on. There is also some pressure for guidelines to be written in a multidisciplinary way. The principal reason, is that the scientific evidence for the effectiveness of the intervention of other health professionals is even more shaky than the evidence for much of what doctors do. For example, there is no evidence that language therapy by trained professionals alters verbal communication after stroke more than support provided by volunteers. The very diversity of physiotherapeutic interventions for common disorders such as cervicobrachial syndromes suggest that there is little evidence of the effectiveness of any. Leaving aside this point about the scientific evidence of effectiveness, there is a danger that if all interventions by all health professionals in one episode of illness are included in one giant guideline, then the whole affair will be too diffuse to have any impact. For major disease areas such as stroke, it might make more sense to have teams of different health professionals working independently on guidelines within their own disciplines, bringing together the results for fused publication, but for mutual critical appraisal.

(12) To get an advisory team together, to undertake a structured critical appraisal of the literature, to consult with users of services, to
consult with purchasers, all takes a considerable time. No sooner is a final package prepared than some other intervention may well be shown to be effective. There therefore has to be a mechanism whereby guidelines can be continually updated, but such updates are few and far between.31

(13) Even in a small country such as the UK, there is surprisingly poor coordination between groups writing guidelines on similar topics. For example, the Royal College of General Practitioners updated its guidelines for the care of patient with diabetes (1985) and asthma (1986) in 1993, without reference to guidelines produced by the Royal College of Physicians in association with the British Thoracic Society31 and the British Diabetic Association32—but these in turn did not refer to the earlier general practitioner guidelines, even though some members were common to each steering group. A particularly striking recent example is that the Nursing Directorate of the NHS Executive has recently commissioned guidelines on the management of leg ulcers from the University of Liverpool, apparently unaware that the Executive’s Health Care Directorate is supporting (through the Royal College of Physicians) the British Association of Dermatologists to write such guidelines; and on the management of hospital acquired infection, although the Health Care Directorate is supporting the college and the Hospital Infection Society in the same endeavour (NHS Executive; paper presented at clinical outcomes group subgroup on clinical guidelines, 29 September 1994).

A proposal for funds to support a ‘guidelines clearing house’ was turned down more than two years ago.

(14) Guidelines can only guide the clinical management of patients already in different parts of the health care system. They can have no impact, for example, on the hospital care of those inappropriately not referred. There is some evidence that suggests that general practitioners refer in terms of the burden that they perceive the local hospital can cope with, rather than according to clinical need.33 Access to care may be endangered by the division between fundholding and non-fundholding general practices.

(15) A further NHS executive letter on ‘Improving the effectiveness of the NHS’ suggested that ‘Individuals and NHS organisations should find nationally produced clinical guidelines ... of assistance in developing local guidelines and taking forward local initiatives to improve clinical effectiveness’.34 There is, however, tension between the need to guide practice using the most robust scientific evidence and the realities of local clinical practice. Local geography, the local availability of skilled staff, and local purchasing priorities (hopefully reflecting the priorities of the local population) must all determine the pattern of service provision. There is also evidence that local development of guidelines aids the chances of their uptake,20 perhaps by increasing a sense of responsible ‘ownership’. However, there is a danger that local priorities in allocating resources might weaken the influence of national guidelines. For example, a national guideline might require that a patient with status epilepticus should be managed in an intensive care unit, based upon evidence of reduced mortality. Local priorities might allocate intensive care beds in another way, or insufficient beds may be provided so that the guideline is downgraded to a statement that such patients ‘required careful monitoring’. Participants at many of the guidelines workshops that I have organised have polarised into two groups—those who propose the best care, and those who propose the best care that is likely to be realisable within the NHS.

**Conclusion**

The vast majority of us in clinical and academic medicine welcome the current international emphasis in delivering good health and health care based upon the best available scientific evidence. There are, however, enormous differences between research studies and the messy realities of everyday clinical practice. The difficulties of translating research into practice lie not only in transferring information, but also in building different systems. Clinicians already recognise the strengths of good research that, potentially, will result in better health. We need to build on the secure foundations of good biomedical and health services research, working as fast as we can towards better everyday clinical practice, while continuing to recognise its complexities.

Some reservations about clinical guidelines


34 NHS Executive. Improving the effectiveness of the NHS. Executive Letter (94)74.