Guidelines, enthusiasms, uncertainty, and the limits to purchasing

Martin McKee, Aileen Clarke

Recently government ministers have set out their vision of the future of purchasing. Ineffective treatments will be discarded and purchasing will be based on guidelines or protocols rather than activity. But have the advocates of this approach considered all the issues? This paper examines the challenges of balancing the desire for protocol based uniformity with the needs of individual patients, explores the extent to which existing purchasing structures can support this process, and questions whether such moves will actually lead to reduced costs. In each case it is concluded that oversimplistic analyses are likely to be misleading and that much of the current debate fails to recognise the complexity of health care.

A health service in which decisions are rational and based on evidence has obvious benefits. Some current treatments are ineffective and others that are effective are underused. Much effort is being devoted to developing clinical guidelines and protocols. Amid continuing disillusion with audit some politicians, managers, and health service researchers see the purchasing process as the way forward. There have been several initiatives to increase the effectiveness of health care through the purchasing process, described in recent executive letters,[1-3] including the "Effective Healthcare" bulletins, the "Outcomes Clearing House," and the changing focus of the research and development programme. These seek to gather together the best available evidence. But it is less clear how this information should be translated into practice.[4]

This vision sees all interventions progressively being evaluated by randomised controlled trials and purchasers undertaking comprehensive assessments of need that will indicate which services should be purchased: locally adapted guidelines will be implemented by all providers; purchasers will buy not activity but guidelines or protocols[5]; change will come about through the "incredibly powerful lever of contracting"[6]; and Dr Mawhinney's seven steps to better purchasing will have been achieved.[7]

But not everyone welcomes the imposition of guidelines through purchasing. Some see it as a political move to reduce professional power, which is an obstacle to the working of the market.[8] Others see it as a crude means of reducing costs, as ministers, citing evidence of widespread variations, argue that much care is ineffective.[9] Finally, some people oppose "cookbook medicine,"[4] in which overrigorous application of guidelines does not permit sufficient recognition of the needs of individual patients. Have those who advocate using purchasing as a means to achieving completely rational, systematic, and evidence based health care really considered how feasible this is? Three questions require answers. What is the role of clinical judgment in the face of inadequate research evidence and legitimate physician and patient preferences? Are there constraints on the extent to which purchasing can be used as a mechanism to improve clinical practice? Can guidelines really reduce costs?

Health Services Research Unit, London School of Hygiene and Tropical Medicine, London WC1E 7HT
Martin McKee, *senior lecturer*
Aileen Clarke, *senior lecturer*

Correspondence to:
Dr McKee.

BMJ 1995;310:101-4

Overcoming clinical uncertainty

The most enthusiastic advocates for the purchasing of guidelines and protocols may have paid insufficient attention to the uncertainty inherent in clinical practice, with the imposition of a spurious rationality on a sometimes inherently irrational process. The purchasing of guidelines or protocols is seen as bringing order into the chaos of widespread variations. Purchasers will stop paying for care that is ineffective. This may be easy when the research evidence of ineffectiveness is clear cut, when a test has low predictive power and adds nothing to other diagnostic information, or when a treatment brings no improvement in health. The history of medicine is full of such examples of treatments once popular but now known to be valueless[10]—for example, gastric freezing or treatments for "night time starvation" or the "wandering womb."

It is also possible, though difficult, to stop purchasing one treatment or diagnostic technique when two are of equal effectiveness but different cost; when less effective treatment is more expensive[11]; or when a treatment is being offered in circumstances in which the outcome is known to be poor, such as a hospital performing a small number of certain procedures.[12] Unfortunately, real life is even more complicated. Most decisions are affected by many different factors, including characteristics of the doctor and the patient and their interaction.

One problem is that treatment guidelines or protocols are useful only once an accurate diagnosis has been made. This depends not only on the clinical ability of the doctor but also critically on the health beliefs and illness behaviour of the patient and the quality of the doctor-patient relationship.

Diagnosis is also subject to bias because of availability error, in which the probability of a diagnosis seems more likely if a doctor has just seen a series of patients with a particular disease and then sees another with similar symptoms[13] or a patient who knows someone with a set of symptoms of a particular serious disease is concerned that he or she has it too. This has stimulated the development of diagnostic guidelines, including computer based expert systems. Unfortunately, these also have their problems.[14] It was recently noted that "diagnostic computer programs have come a long way, but they still have a long, long way to go."[15]

Once a diagnosis has been made, factors affecting the clinical decision, and thus the role that treatment guidelines or protocols can play, fall into three categories. These are the probability of particular outcomes, the valuation of the different outcomes that may result, and willingness to accept or live with a

degree of risk. Only the first of these is addressed directly by the effectiveness agenda.

Estimates of the risks and benefits associated with various courses of action are often inaccurate[16] but research based guidelines rarely provide all the answers. Panels examining appropriateness, using formal consensus methods, frequently identify many indications where there is uncertainty, partly because of other, often unquantifiable factors.[17]

A further problem is that outcomes can be expressed prospectively only as probabilities applying to populations. Consequently, studies cannot predict with certainty the outcome for an individual. Though surgery for glue ear is often unnecessary, and is more likely to be ineffective if the hearing loss is less than 25 dB, some children with lower levels of hearing loss do benefit.[18] Prostatectomy is much less likely to relieve mild than severe symptoms but, again, some men with mild symptoms benefit.[19]

VALUING OUTCOMES

Valuation of outcomes and dealing with risk are still more difficult. Even if the probability of a particular outcome can be ascribed to individual patients with reasonable certainty, each outcome may be valued differently. Two 40 year old pregnant women facing the decision about amniocentesis and, if the fetus has Down's syndrome, subsequent termination of pregnancy may be in possession of accurate information on the risks of all of the possible outcomes but may make different decisions because of the differing values that they place on each outcome. This is true of many clinical situations.[20,21] When considering diagnostic tests, individuals place different values on information. This may be impossible to predict as it may be dependent on the result. In one study patients with possible multiple sclerosis who were found to have the disease derived benefit from possession of the information in terms of quality of life whereas those who remained without a firm diagnosis were worse off.[22]

Economists have developed methods of measuring utilities ascribed by patients to particular outcomes, but though it is possible to obtain aggregate values for groups of people, individuals' values vary widely, both with regard to the value placed on particular situations and with regard to the timing at which those states occur. Utilities, particularly at the point of accepting or rejecting an intervention, "cannot be averaged across individuals,"[23] Also the result obtained is sensitive to the method used and the questions asked.[24]

REX FEATURES

A health service in which decisions are rational and based on evidence has obvious advantages

Finally, even with accurate knowledge of risk and agreement on the valuation of outcomes there may be differing views on what constitutes an acceptable risk. People tend to be more averse to taking a risk when a potential gain is involved but become risk takers when a loss is possible,[25] and specialists and general practitioners have different time horizons when assessing risk.[26] However, other factors, many of them individual, are important, such as previous experience of treatment.[27]

Research based protocols and guidelines can identify the probability of various outcomes, facilitating informed choice,[28] and can also stimulate debate on the value of each outcome. They cannot include individuals' values of different outcomes or change their willingness to accept risk.

Guidelines should therefore only ever by used as guidelines. Purchasers seeking to convert them into restrictive protocols in contracts will have much to do to incorporate patient choice in a way that is distinguishable from the unjustifiable variation in clinical judgment that they seek to combat. This raises the question of monitoring. If purchsers are to move from purchasing activity to purchasing protocols,[4] then accountability for public money requires them to decide how to measure what they are buying.[3] This may be possible in an all or nothing situation—for example, stopping diagnostic dilatation and curettage in all women under 40[29]—even allowing for the limited discriminant power of routine coding systems[30] and the opportunities for gaming through shifts in definitions of diagnoses and other variables.[31] But it is much more difficult when there is legitimate variation. It is for these reasons that Wennberg concluded that "micromanagement" of the doctor-patient relationship is impossible.[32]

Purchasing mechanism

Even if the view is accepted that guidelines should not be imposed too rigorously and they are used to educate rather than compel, current purchasing structures provide insufficient support for this process. On the one hand, there is recognition that the purchasing function is underdeveloped.[33-35] On the other, those trying to develop purchasing are faced with continuing environmental turbulence, including mergers of health authorities, mergers with family health services authorities, mergers with subsequent abolition of regions, and, most recently, substantial increases in the coverage of general practice fundholding.[36] At the same time the Department of Health is seeking to reduce management costs further in a system that by its nature requires greater investment in management than the one it is replacing, if there is to be a sustained attempt to promote effective and equitable care.

Many would argue that the purchasing process can never make a major contribution to increasing the effectiveness of health care. Some of the most compelling reasons for this view include, firstly, the built in flaws in the functioning of any internal market in health care and the scope for opportunism that this affords.[37] Secondly, the quantity and extent of "quality police" who would need to be employed to ensure adequate continuing external quality assessment[38]; and, thirdly, the fact that most available evidence suggests that locally acceptable activities perceived as internal to an organisation or a group of clinicians (for example, education by local opinion leaders) are the most effective at changing practice.[39]

Many staff in purchasing organisations have little formal training in the concepts of effectiveness and appropriateness, though groups in at least two English regions are actively seeking to rectify this. Contracts

REX FEATURES

There may be many people in the community who have symptoms but have not sought treatment. Hence implementing guidelines through purchasing may increase the volume of activity and thus costs

managers often deal with a wide variety of contracts covering a multiplicity of specialties. Medical mystification can be used extremely effectively by clinicians seeking to diminish the ability of sometimes rather junior contracts mangers to have an impact on their practice. Public health can make an important contribution, but attempting to keep up to date with the minutiae of change in clinical practice in several different specialties may not represent the most efficient use of public health resources. Purchasing by fundholding general practitioners may overcome some of these problems but introduces others, particularly in relation to diseconomies of scale, transaction costs, or the relative powerlessness of small fundholders in the face of major providers.

Saving costs

Purchasing can be viewed as a means of either saving costs or increasing the effectiveness of care. For many people it is seen as a combination, though with differing degrees of emphasis on one or the other. Arguments for using guidelines as a means of reducing costs are based on the awareness of extensive variation in health care intervention rates. This variation exists at all levels of the health care system and has traditionally been ascribed to clinical uncertainty and, more recently, to clinical enthusiasm,[40] doctors doing more of what they enjoy and less of what they do not.

There is an implicit belief that high intervention rates can be equated with a degree of ineffective care, a view that underpins the comparative approach to assessing need.[41] But there is little evidence that this is true. Studies of high and low use areas in the United States found similar levels of inappropriate interventions.[42] Those who intervened more did so equally on patients with appropriate and inappropriate indications. Guidelines can be seen as a means of reducing the intervention rate of the more enthusiastic doctors, a view supported by most published studies that have focused on means of reducing unnecessary investigations.[43]

The view that the purchasing of guidelines should be used to increase the amount of effective care and that this is congruent with reducing costs stems from this view. Guidelines may, however, increase the volume of activity and thus cost. There is growing evidence of a large number of people in the community with symptoms that could be relieved effectively but who have not sought treatment.[44,45] Calculations of costs and benefits should include the cost of developing, implementing, and monitoring the guidelines. Experience with medical audit[46] in the United Kingdom and professional review or the work of the Agency for Health Care Policy and Research in the United States suggests that this may be considerable.

The work entailed in developing and implementing guidelines, then, should not be underestimated. Guidelines need to be developed locally and disseminated within the context of an educational programme.[37,47] These processes take time, effort, skills, and money. And the continuing lack of appropriate research based evidence in many areas is a major problem. Poorly developed or insufficiently comprehensive protocols or guidelines can be inappropriate, unhelpful, or even potentially misleading.[48]

The way forward

At one level the issue of implementing guidelines through purchasing reflects a wider debate about the validity of applying models based on rational decisions and perfect information to complex adaptive human systems.[49] Dogmatic approaches based on these models seek to avoid market failure by increasing the quality of information used to inform transactions, though the advocates of these policies often ignore the risk that in doing so they may precipitate market failure through increased administrative costs or failure to value such qualities as altruism. These are fundamental issues that need to be addressed explicitly by health policy makers, though they may be seen as somewhat esoteric by those charged with implementing government policies.

With respect to more practical issues, this review has highlighted some limitations of the purchasing process as a means of increasing the effectiveness of health care that is delivered. This should not be interpreted as reflecting either complacency that all is well or that nothing can be done. Purchasing clearly has a contribution to make, though it may be more limited than originally anticipated. Large areas of clinical practice will remain where attempts to purchase by means of protocols will be impossible. In these areas more appropriate approaches including increased reliance on education and clinical audit to improve quality of care are likely to offer the best way forward.[37]

Implementation must be sensitive to the concerns of both health care professionals and their patients. There must be recognition that cooperation achieves better results than conflict and coercion.[50] Given the importance of the "doctor as medicine," forcing doctors to abide by guidelines that they do not believe in—especially if this is obvious to the patient—is likely to have an adverse effect on outcome. Ways of reconciling patient choice with research based evidence are needed when they conflict.

Perhaps the most difficult issue for all of us concerned in purchasing or providing care is the need for an increased acceptance that a degree of uncertainty is integral to health services. When Heisenberg introduced the uncertainty principle over 60 years ago,[51] this did not halt progress in theoretical physics. Rather, by explicitly identifying the limitations to knowledge it enabled scientists to pursue the achievable rather than the impossible.

Certainly a more explicit recognition of the limits to and costs of implementing guidelines and protocols through purchasing may inject a greater degree of reality into some of the current debates.

We are grateful to Colin Sanderson, Naomi Fulop, Nick Black, Nick Mays, and Jenny Roberts for helpful comments.

1 NHS Executive. *Improving clinical effectiveness.* Leeds: NHSE, 1993. (EL(93)115.)
2 NHS Executive. *Improving the effectiveness of the NHS.* Leeds: NHSE, 1994. (EL(94)74.)

3 NHS Executive. *Priorities and planning guidance for the NHS: 1995/96.* Leeds: NHSE, 1994.
4 Hayward J. Purchasing clinically effective care. *BMJ* 1994;309:823-4.
5 Sheldon TA, Borowitz M. Changing the measure of quality in the NHS: from purchasing activity to purchasing protocols. *Quality Health Care* 1993;2: 149-50.
6 Delamothe T. Wanted: guidelines that doctors will follow. *BMJ* 1993; 307:218.
7 Mawhinney B. The vision for purchasing. Quoted in: *Purchasing for health. A framework for action.* Leeds: NHSE Health Publications Unit, 1993.
8 Bunker J. Can professionalism survive in the marketplace? *BMJ* 1994;308: 1179-80.
9 Extracts from UK secretary of state for health's speech to BMA/King's Fund/ Patients' Association conference, 11 March 1993. *Health Care Analysis* 1994;2:5-12.
10 Beeson P. Changes in medical therapy. *Medicine* 1980;59:79-84.
11 Song F, Freemantle N, Sheldon T, House A, Watson P, Long A, *et al.* Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability. *BMJ* 1993;306:683-7.
12 Luft HS. The relationship between surgical volume and mortality: an exploration of causal factors and alternative models. *Med Care* 1980;18: 940-59.
13 Sutherland S. *Irrationality.* Harmondsworth: Penguin, 1994.
14 Berner ES, Webster GD, Shugerman AA, Jackson JR, Algina J, Baker AL, *et al.* Performance of four computer-based diagnostic systems. *N Engl J Med* 1994;330:1792-6.
15 Kassirer JP. A report card on the computer-assisted diagnosis—the grade C. *N Engl J Med* 1994;330:1824-5.
16 Lictenstein S, Slovic P, Fischhoff B, Layman M, Combs B. Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory* 1978;4:551-78.
17 Hunter D, McKee M, Sanderson C, Black N. Appropriate indications for prostatectomy—results of a consensus panel. *J Epidemiol Community Health* 1994;48:58-64.
18 Black NA, Sanderson CF, Freeland AM, Vessey MP. A randomised controlled trial of surgery for glue ear. *BMJ* 1990;300:1551-6.
19 Flood AB, Black NA, McPherson K, Smith J, Williams G. Assessing symptom improvement after elective prostatectomy for benign prostatic hypertrophy. *Arch Intern Med* 1992;152:1507-12.
20 Kassirer JP. Incorporating patients' preferences into medical decisions. *N Engl J Med* 1994;330:1895-6.
21 Barry MJ, Mulley AG, Fowler FJ, Wennberg JE. Watchful waiting versus immediate transurethral resection for symptomatic prostatism: the importance of patients' preferences. *JAMA* 1988;259:3010-7.
22 Mushlin AI, Mooney C, Grow V, Phelps CE. The value of diagnostic information to patients with suspected multiple sclerosis. *Arch Neurol* 1994;51:67-72.
23 Verkes RJ, Thung PJ. Medical decision analysis and the coming moral crisis in health care. In: Jensen UJ, Mooney G, eds. *Changing values in medical and health care decision making.* Chichester: John Wiley, 1990:81-95.
24 Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use.* Oxford: Oxford Medical Publications, 1991.
25 Kahneman D, Tversky A. Choices, values and frames. *Am Psychol* 1984;39: 341-50.
26 Otten R, Widdershoven G. Social change and the perception of risks in medical practice. In: Jensen UJ, Mooney G, eds. *Changing values in medical and health care decision making.* Chichester: John Wiley, 1990.
27 Stiggelbout AM, Kiebert GM, Kievit J, Leer JW, Stoter G, De Haes JC. Utility assessment in cancer patients: adjustment of the time trade-off scores for the utility of life and comparison with standard gamble scores. *Med Decis Making* 1994;14:82-90.
28 Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I. Developing scientifically valid guidelines. *Quality Health Care* 1993;2: 243-8.
29 Coulter A, Klassen A, McKenzie I, McPherson K. Diagnostic dilatation and curettage: is it used appropriately? *BMJ* 1993;306:236-9.
30 Iezzoni LI. Using administrative diagnostic data to assess the quality of hospital care: pitfalls and potential of ICD-9CM. *Int J Technol Assess Health Care* 1990;6:272-81.
31 Assaf AR, Lapane KL, McKenney JL, Carelton RA. Possible influence of the prospective payment system on the assignment of discharge diagnoses for coronary heart disease. *N Engl J Med* 1993;329:931-5.
32 Wennberg JE. Outcomes research, cost containment, and the fear of health care rationing. *N Engl J Med* 1990;323:1202-4.
33 Organisation for Economic Cooperation and Development. *OECD economic surveys 1993-1994.* Paris: OECD, 1994.
34 Gill M. Purchasing for quality: still in the starting blocks? *Quality health care* 1993;2:179-82.
35 Kirkup B, Donaldson LJ. Is health care a commodity: how will purchasing improve the National Health Service? *J Public Health Med* 1994;16:256-62.
36 NHS Executive. *Developing NHS purchasing and GP fundholding.* Leeds: NHSE, 1994. (EL(94)79.)
37 Roberts JA. Managing markets. *J Public Health Med* 1993;15:305-10.
38 Boufford JI. US and UK health care reforms: reflections on quality. *Quality in Health Care* 1993;2:249-52.
39 Haines A, Jones R. Implementing findings of research. *BMJ* 1994;308:148-92.
40 Chassin MR. Explaining geographic variations: the enthusiasm hypothesis. *Med Care* 1993;31(suppl):YS37-44.
41 Stevens A, Raftery J, eds. *Health care needs assessment.* Vol 1. Oxford: Radcliffe Infirmary, 1994.
42 Chassin MR, Kosecoff J, Park RE, Winslow M, Kahn KL, Merrick NJ, *et al.* Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA* 1987;258:2533-7.
43 Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;342:1317-22.
44 Hunter D, McKee M, Sanderson C, Black N. Prevalence and severity of urinary symptoms in British men aged 55 and over. *J Epidemiol Community Health* (in press).
45 Harries U, Landes R, Popay J. Visual disability among older people: a case study in assessing needs and examining services. *J Public Health Med* 1994;16:211-8.
46 McKee M. Is money wasted on audit? *J R Soc Med* 1994;87:52-5.
47 Grimshaw JM, Russell IT. Achieving health gain through clinical guidelines. II. Ensuring guidelines change medical practice. *Quality Health Care* 1994;3:45-52.
48 Gabbay J, Stevens A. Towards investing in health gain. *BMJ* 1994;308: 1117-8.
49 Gell-Mann M. *The quark and the jaguar.* London: Little, Brown, 1994.
50 Axelrod R. *The evolution of co-operation.* Harmondsworth: Penguin, 1990.
51 Heisenberg W. *Physical principles of the quantum theory.* Chicago: Dover Press, 1930.

(Accepted 8 November 1994)

Management for Doctors

Conflict, power, negotiation

Liam Donaldson

This is the fourth in a series of articles dealing with issues arising as clinical practitioners increasingly take on managerial roles. The series is edited by Jenny Simpson

Northern and Yorkshire Regional Health Authority, Newcastle upon Tyne NE6 4PY
Liam Donaldson, *regional general manager and director of public health*

BMJ 1995;310:104-7

Conflict, how it arises and how it is resolved, is closely related to where power lies within a hospital, health authority, or medical practice and how this is influenced by external factors. This is a subtle and complex process which differs enormously from one health care organisation to another, over time, and from issue to issue. Power can influence change in predictable and conventional ways. For example, in the choice of a new member of the medical staff of a hospital, the head of the relevant clinical department and his or her consultant colleagues are likely to be the principal determinants of the type of person chosen. On the other hand, a major and unforeseen impact on the organisation might be produced by a quite junior member of staff whose power derives from the possession of information which could be passed anonymously to local media and arouse public concern.

Conflict exists whenever individual or group interests diverge within an organisation, and if its values or goals are at odds with those of the external environment. These considerations apply to the operation of both private and public sector enterprises, but in the public sector the influence of the wider public and political dimension makes the process of setting and achieving objectives in an orderly way much more complex. In health care organisations the potential for conflict arising through internal and external factors is always present. The resolution of such conflict is often the route to progress or the way in which major change takes place. It is one of the jobs of management to understand the potential sources of conflict and to be able to predict how, when, and why they will arise. Similarly, effective management of change is not possible without a clear understanding of the sources of power within the organisation and how they can be harnessed, not just to resolve conflict, but to bring about improvement and generate innovation.

The NHS reforms in 1990 introduced new mechanisms for the organisation and funding of health care in Britain.[1] By separating the responsibility for purchasing care from that for its provision, these reforms sought to reorientate the management of the service by changing the role of existing organisations. The functioning of the reorganised health service now depends on the interaction between bodies that purchase health care for populations (health