
Book Section:
Culyer, A. J. (Anthony J.) (2012) Chisels or screwdrivers? A critique of the NERA proposals for the reform of the NHS. In: The humble economist : Tony Culyer on health, health care and social decision making. University of York , York , pp. 23-37. ISBN 9780952560159

---

2 CHISELS OR SCREWDRIVERS? A CRITIQUE OF THE NERA PROPOSALS FOR THE REFORM OF THE NHS

by Professor A J Culyer

Introduction

The report, *Financing Health Care*, prepared by National Economic Research Associates, (NERA) runs to two volumes and 1453 pages, reviews the experience of 12 countries, develops a Prototype model and a phased strategy by which the UK might approach it, and has plainly involved a huge amount of work to which I can scarcely do justice in the space available. I thus have a problem akin to that of the authors of the report: given the vast complexity of the systems – indeed of our own – and the immense range of issues that any proposals for change must range over, how can one fasten on to the key issues without grossly over-simplifying and doing gross injustice to the authors? The problem faced by the authors – which I may as well say that I do not think they have altogether resolved – is that, in comparing international systems, while it is undoubtedly true that some aspirations and problems are shared in common, history and culture never are, and neither are *all* of the aspirations and problems. It is not so hard (though it is not easy) to compare the workings of one system with either some aspirations of one's own or those of another system and find it wanting – but that is not a lot of help if the shortcomings thus identified are being judged by a set of criteria which are not those of the clients of the system being evaluated.

I propose to divide my discussion into the following parts:

- first I shall review the Report's *diagnosis* of the problems faced by the NHS;
- second, I shall review their proposed *treatment*;
- thirdly, I shall then ask whether the diagnosis is *correct* and the treatment *appropriate* and *cost-effective*;
- finally, I shall briefly enquire into other possible diagnoses and alternative treatments for the patient.

I shall not review the descriptions and analyses of the health care systems of other countries but shall focus on the UK, making occasional reference to other countries where it may be helpful. I shall not comment on the study's discussion of the implications of its recommendations for the pharmaceutical industry, though I may as well say, less there by any doubt about it, that a profitable, thriving and innovative pharmaceutical industry is something to cherish and, although I do not think that the welfare of the industry is something that should drive the financial and organisational structure of health services, I am confident that the industry is well capable of responding effectively, appropriately – and profitably – to most systems likely to evolve in western society that are this side of rational.

The diagnosis

There is no single place in the study at which the reader can find a convenient summary of the problems faced and so there is some risk that I may have missed some crucial element in the diagnosis, However, a trawl reveals the following:

- government is predicted to relinquish the roles of health insurer and health care provider to the market, which forces the issue of change and reform dramatically (p3);
- the demand for health care will continue to outstrip the supply due to aging population and technical advance (p3);
- bureaucracies are inherently less responsive to demands for new treatments than market orientated systems (p22);
- governments will choose to spend less on health care than individuals prefer (p3);
- centralised systems, such as the NHS, create distortions which seriously compromise the delivery of appropriate levels of service (p5);
- patients and physicians do not have the correct incentives or information for making well-informed and efficient choices (p6).

The treatment

The recommended treatment is 'required to move the existing UK system towards that of the NERA Prototype over the long term' (p1127). The main components are:

- establish an agreed *Guaranteed Health Care Package* (GHCP) (the initial contents of this are said to be the current range of NHS services, whatever this is (p1132), but the 'target' contents of the GHCP in the Prototype are not given, so it is not clear whether the services included would curtail or extend the current range of NHS and community care benefits) (p1127);
- introduce a market for health care insurance for the GHCP (plus top-ups at customer discretion) with insurers being denied the right to turn away clients for the GHCP at prevailing premiums (p1127-8);

- make health insurance compulsory (p1128);
- set premiums in two parts: one, a function of (family?) income payable to a central agency, for redistribution to insurers after adjustment for risk: the other payable directly to insurers and risk-rated (p1129);
- establish a public National Health Insurance Fund (NHF) for those 'unable to manage their own health care insurance' (p1130);
- all services within the GHCP to be subject to mandatory co-payments (initially with exemptions but 'increasing the rigour over time') (p1133);
- by implication, abolish the NHSE, its regional offices, FHSAs, DHAs and GP fundholders (it is not clear what implications there are for services provided outside general practice or hospitals, such as community services, blood transfusion services or ambulance services);
- make contracts between insurers (purchasers) and providers (public or private, primary, secondary and tertiary care) legally binding and enforceable at law (p1131);
- deny insurers (purchasers) the right to own providing institutions (p1131);
- create unregulated entry for providers (subject to 'medical qualification requirements') and permit providers full access to the capital market (p1131);
- reduce the role of government to accrediting insurer, enforcing compulsion in insurance and the way the insurance market works, collecting premiums for the central fund and specifying the GHCP.

Is the diagnosis correct?

What are the objectives?

In asking whether the diagnosis is right, one comes directly up against an issue to which I alluded at the beginning: what are we trying to achieve? It is quite clearly one central objective of the current NHS to increase the scope and range of individual (patient) choice. I shall return to this later. Another, on which I shall for the present concentrate, is to maximise 'health gain'. What this means is not entirely clear and neither 'health' nor 'gain' (presumably some positive difference attributable to the use compared with the non-use of health services) are easy either to conceptualise or quantify. However, these difficulties are hardly grounds for ignoring or replacing this objective with some other. Indeed, if I were going to write a report on the current problems of the NHS I would actually begin with an analysis of the *efficiency* issue of what it ought to be maximising (and what limits its success in accomplishing that objective) and then complement that by a parallel analysis of the currently unpopular theme of *equity* in resource distribution (and what limits success in accomplishing that). I think I might be able to make a good case for maximising health gain (and justify this broad objective as superior to a consumer sovereignty model), and be able to put some practical content into both 'health' and 'gain' (with perhaps some epidemiological help) to guide both purchasers and providers, and I think also that I might be able to develop both some principles to guide distributional judgements and some practical suggestions for improvements on where we are now. I do not think the implications of this for organisational and financial structures would be terribly radical – but, then, I take the view that the structures currently being developed (which may be characterised as *'demand-side socialism'*) are broadly right with the main things needed being a loosening of the capital market, some mechanism for freeing management from political interference, and a need for equity to take a more central role as an allocation criterion – plus one other thing (where I come into closer tangency with the authors of our report) – the need to create a mechanism through which genuine desires by the purchasing/voting public for greater expenditure on health care can be reflected in the actual

resource flow to the NHS, thus escaping an implacable Treasury public expenditure constraint imposed on grounds of so-called macro efficiency but at the cost of probable micro inefficiency. But this is not *my* report. I make these comments simply to remind you of the old truth that where you stand (on health care reform) depends on where you sit (your ideological and financial interests) and where you sit is essentially a matter of the objectives you either accept from someone else or put up for your little old self.

So what are the underlying objectives for the health care system of the UK as seen by our authors? They are elusive. And who are the ultimate clients for the study and what are their objectives? Are their objectives likely to be shared by the British electorate? Even if we pick the diagnosis apart, these issues do not become any clearer. But let us look anyway at the diagnosis in more detail.

The diagnosis in detail

(i) A reduced role for Government

I do not intend to give any weight to the first of their diagnostic bullet points because, as a prediction, I see no basis for it other than as wish fulfilment. The issue as I see it is whether the government ought to relinquish or take on roles, not that this is something to be taken for granted with the implication that we then have to cast around to make the best out of whatever fate thrusts upon us.

(ii) Demand for health care will outstrip supply

The second bullet point is more substantive. But it is hardly news. What it is saying is that health care has to be rationed. I don't think anyone denies that (apart from the occasional minister who wants to avoid an awkward public debate). The issues are, of course, the *levels* to which demand (or need) is to be rationed, the *criteria* that are to be used in the rationing and the *means* used to do the rationing. But let me enter a dissenting note of caution on those alleged twin drivers of the overall medical bill, to which the authors draw our attention: an aging in population and changing technology. I find it distressing that our authors, along with a good deal of distinguished company, treat these two factors as *exogenous* when it seems to me plain that they are in very large part *endogenous* – that is, determined within the overall economic system. If health care expenditure per head of elderly is rising relative to health care spending per head of the rest, then that is the result of decisions taken within the system – and decisions, moreover, that are frequently alleged, by people in a position to know, to be inappropriate. Chucking high-tech medicines and inpatient care at the elderly regardless of true cost-effectiveness is not something that we *have* to do. Nor do we have to adopt every latest mark of imaging technology the moment it appears (indeed, the authors themselves make quite clear that different systems manage to control the introduction of new technologies at quite different rates). Incidentally, are such technologies to be in or out of the GHCP? You could argue either way – or for a sophisticated mix. But to determine which one needs a clearer statement of the objectives of the system.

Moreover, I conjecture that the character of the research that produces the sorts of technology that are held to drive costs ever upwards (relative to constant price GDP) *is itself endogenous* – it itself is driven by knowledge of what it is that the finance of medical care will pay for, so even the research (and especially the industrial research in companies and the research sponsored by them in universities and elsewhere) is ultimately endogenous and therefore influenced by system design. Of course, some technological change is cost-reducing rather than cost-increasing. The development of an effective vaccine for Polio is a classic example that eliminated the need for the iron lung; or that for rubella, which led to a greatly reduced incidence of babies with birth defects. But, in general, technological advance in medicine tends to be cost-increasing. A notable example is modern neo-natal intensive care which has increased the survival chances greatly of low and very low birthweight babies but which has major cost consequences not only of the neonatal care itself but also of the subsequent long term care of these children as they survive into adulthood. I conjecture that this 'bias' towards cost-increasing technological change is not accidental. None of these things is inexorable. They are themselves generated by the systems we have and the incentives they embody. No successful business is going to embark on the development of products if

it believes there is no market for them – and whether or not there is such a market depends on the willingness and ability to pay of those with power to decide what technologies they shall use, and the criteria to be used in selecting new – or come to that, old – technologies. These ultimate determinants of the pattern of technological research and development are all endogenous and therefore a function of system design.

(iii) Bureaucracy is less responsive to new technology

It is not very good analysis simply to say that 'bureaucracies are inherently less responsive to demands for new treatments than are market-orientated systems' (p22). Market orientated systems of competing funders have *immense* bureaucracies of their own which respond, as do all bureaucracies, to the organisational goals that are set for them and the rewards systems in operation to promote those goals. It *may* be that public bureaucracies are less efficient than private ones – but we have to ask (again): *efficient at what?* If a public 'bureaucracy' like a purchasing health authority has more rigorous standards of effectiveness than a private health insurance agency, then the difference between purchasing decisions will reflect something quite different from 'inherent' lack of responsiveness. After all, there is abundant evidence that competition between health care providers in the US operates less through price than through what is, somewhat misleading, often called 'quality', and this is why you will find under-utilised (and probably mis-used) CAT scanners in neighbouring 50-bed hospitals in the US. Is this the sort of 'responsiveness' which our authors want to see in the UK? Perhaps it is. But if it is not, I have to wonder at the (bureaucratic) mechanisms that competitive insurers might employ to counter the very real inherent tendencies that such competition is likely to evoke, especially since it seems inevitable that they will seek to fund 'managed care' on the provision side of health care.

Beneath all this there is, however, a fundamental and real difficulty. It can be put quite simply as a question: what is the optimum rate of diffusion of a new technology? It is not adequate to reply; 'let the market decide' because, first, the market is extremely imperfect and, second, the element of public accountability for expenditures is going to be high even under the prototype. In essence the problem involves a trade-off between two uncertain elements: the postponement of possible (but uncertain) benefits while effectiveness and cost-effectiveness trials and analyses are done, against the greater assurance that what is adopted will have real benefit and constitute value for money. These issues are currently being examined in the Health Technology Assessment programme of the NHS Executive and I do not pretend to know what the right general answer is – except that it is unlikely to be 'leave it to the market'! To whom will our authors leave it?

(iv) Macro and micro efficiency

I conjecture that the nub of our authors' diagnosis lies in the last three bullet points. They are telling us that too little is spent on the NHS (macro inefficiency) and that what is spent is not spent efficiently (micro inefficiency). Now, efficiency is, of itself, a pretty emotionless term. It means simply maximising outcome per unit of input. The big issue here is evidently *what should we be maximising?* I think what our authors have in mind is good old-fashioned welfarism – we should be maximising individual welfares, as perceived by individual clients. Put more crudely, health is like most other things, so let individuals choose subject to the prices they confront and their incomes. It must be said, less I be thought unfair, that it is not all that clear that this really is what is in their minds, particularly when one reads their discussion of performance (p13-15). They tell us (p13) that "'*Health care expenditure (HCE) as a percentage of GDP*', and '*HCE per capita*' are measures for **macroeconomic efficiency** of health care systems' but they also tell us (p29) that 'the amount… of health care services should reflect the informed preference of consumers'. Now, as a matter of fact (or, rather, of definition) HCE as a percentage of anything or per head of population tells us nothing at all (even as a proxy) about efficiency of any kind, nor would a monetary estimate of benefit as a percentage of something or per head. 'Macroeconomic efficiency', if it means anything, must mean that the total spend is such as to optimise the social benefit relative to the social cost. The fact, as reported by our authors (p13), that HCE as a percentage of GDP has been *rising* in

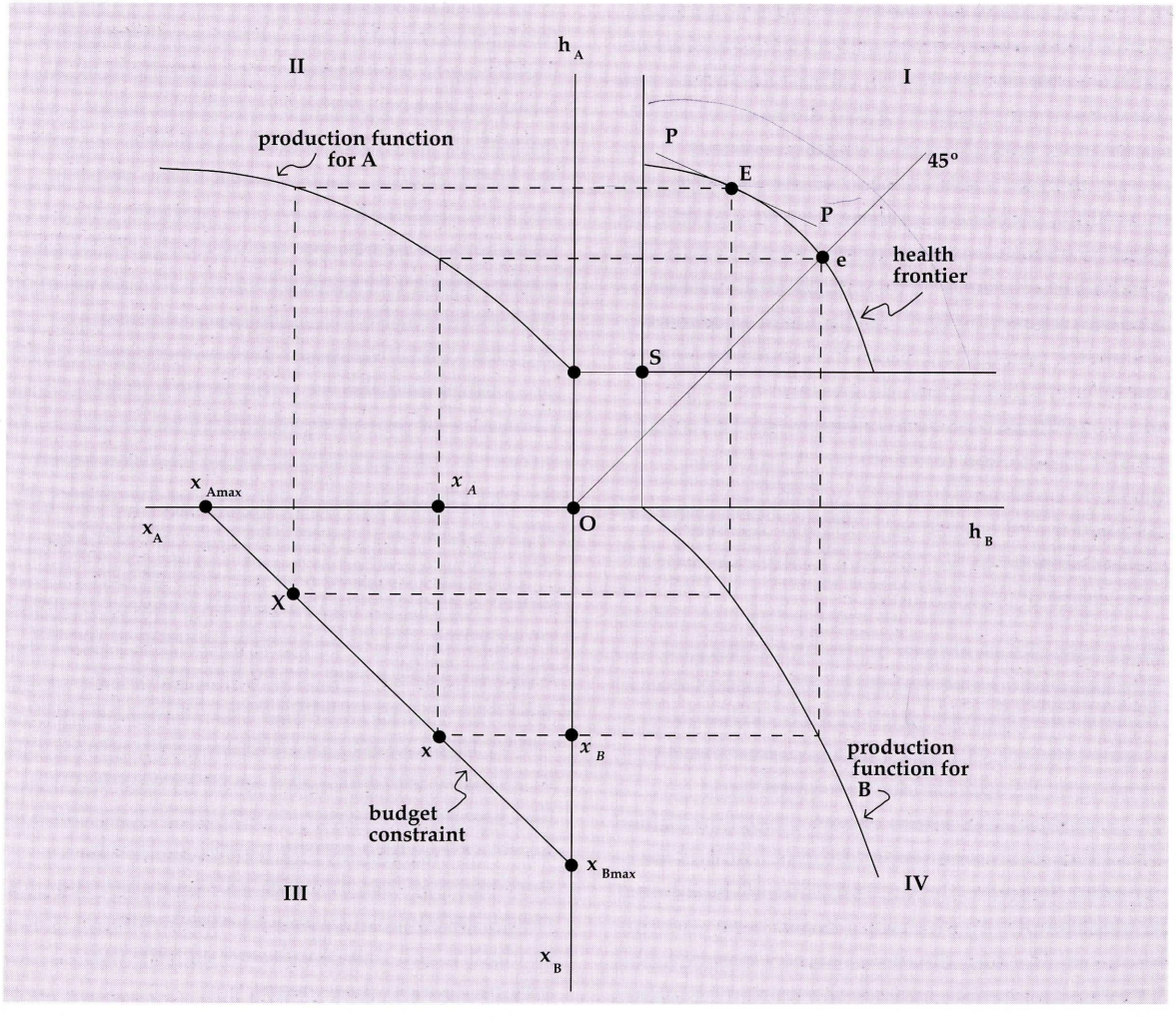
the United States cannot credibly be held to imply that the macroeconomic efficiency of US health care has been rising. Or should we take it that our authors want us to interpret the trend as evidence for a *falling* macroeconomic efficiency in the US? The ambiguity is plain. Likewise, the fact that the UK spends less on health care as a percentage of GDP or per head than many other developed countries, does not imply anything about efficiency unless one assumes that by some magic everyone else has their shares at the optimal level *and* the criteria for determining cost and benefit are appropriately the same across comparator countries.

Our authors' index of *microeconomic efficiency* is physician visits per head. This index is neither a cost nor a benefit indicator, but a measure of utilisation or, if turned upside down, a workload measure for physicians. So this, in international comparisons, is not going to reveal any relative efficiency or inefficiency in the NHS to other countries. Nor does it have any clear bearing on the issue of whether these quantities 'reflect the informed preference of consumers'.

Back to the objective of 'efficiency'

Economics has only one concept of efficiency, not a micro one and a macro one, and central to this concept is the idea of maximising some value function, such as utility, welfare or health gain, subject to constraints. Depending on the scope of the analysis, the constraint might be taken as the resources available to the NHS (where the system is judged efficient if it maximises the postulated value function given these resources) or, at a broader level, the resources of the whole

Figure 5 **Health care efficiency and equity diagram**



community (where the health care system is judged efficient if it allocates an optimal share to health care and maximises the value function given the share allocated to it).

The idea is at once so important and easy to lose sight of when one comes to try to apply it, that it is worth spending just a little time getting the theory straight. I hope you will forgive me for an economist's diagram that I use to illustrate efficiency in the context of a given 'budget' for health care. In Figure 5 I have assumed, for simplicity, that there are just two people (A and B) and that there is a fixed sum annually available for spending on health care. This budget is shown in quadrant III of the diagram as sum of money, which could all be spent on A (in which case A gets X_{Amax}) or on B (X_{Bmax}). Or, of course, it could be divided between them at any point on the straight line connecting these two points, which is called the *budget constraint*. The outcome of applying health care resources to A or B depends upon what economists call the *production function*. A production function shows the *maximum* rates of outcome that can be obtained at various levels of use of the inputs, assuming that at any level of opportunity cost, the mix of inputs (bed days, physician time, use of equipment, etc.) is optimised to maximise expected outcome. These functions identify what the authors quite correctly call (p150) 'productive efficiency'. I have supposed that the production functions exhibit diminishing returns so that, for constant increases in resource commitment under prevailing technology, additional health outcome becomes smaller and smaller. Production functions for A and B are shown in quadrants II and IV, where I have assumed that B is relatively sick and also has a greater capacity to benefit from health care over a wide range of expenditures.

I have also assumed that 'health' is the relevant outcome desired and that we have an acceptable measure of it. Each production function has its origin at the relevant individual's current health status or some appropriately weighted (and discounted) sum of future expected health without a health care intervention, and the outcome is the expected *change* in (the present value of) health that results from the application of health care resources. The co-ordinates of point S in quadrant I, the 'starting point', indicate the presenting states of health of the two individuals.

If we were to trace round the maximum health gain for A and B for different divisions of the budget in quadrant II between them, one would trace out the convex locus in quadrant I which is called the *health frontier*. This shows the maximum increases in health that are possible, given the budget in quadrant III, the presenting states, and the two production functions determined by prevailing technology in quadrants II and IV.

In the sense of production efficiency, *any point on the health frontier is an efficient point*. It follows that a system is inefficient if, for any reason, the allocation of resources between individuals (in quadrant III) results in a point in quadrant I that lies inside the health frontier. This is most likely to arise because the most efficient production technology is not being used so that, given any amount of resources devoted to, say, A, the health gain is less than that indicated by the production function in quadrant II. In common parlance, such treatments would be seen either as inappropriate, inefficacious, ineffective, or not cost-effective. A failure in any one of these respects for either individual will result in an outcome that lies below the health frontier in quadrant I. In my opinion, this source of inefficiency is the most prevalent in *all* systems of health care including the NHS (the epidemiology literature is replete with examples of wide and inexplicable variations in clinical practice, continued use of proven ineffective procedures, and the use – this is the largest category – of procedures that have never been subject to careful critical scrutiny, let alone tested for cost-effectiveness). A concern for this aspect of efficiency lies at the heart of the government's drive for a 'knowledge-based' health service and which is also driving its research and development programme, the meta-analytic function of the UK Cochrane Centre, and the no less important activities of disseminating best practice (the York Centre for Reviews and Dissemination) and trying to secure a contracting and incentive environment that maximises take-up of cost-effective methods of health care, which mainly means educating purchasers so that they can better identify true needs, and not just current mortality and morbidity, and then purchase truly effective services to meet these needs.

There is a big value judgement in the foregoing, which is that I have chosen a particular outcome concept – health. However, I am hopeful that the analysis so far can be tied in quite closely to our authors' analysis because they too place considerable emphasis on health status (proxied negatively by potential life years lost and perinatal mortality) as an objective for systems and as a measure of performance.

Some economists would push the idea of efficiency further – beyond that of *productive efficiency* – so as to embrace the idea of *exchange efficiency* – an efficient allocation of outcome *across individuals*. In the market system, individuals ('consumers' in the language of our authors) would form a judgement of the value of additional health to them and express this in terms of the purchase of appropriate inputs, given their income and wealth, insurance status, and so on. The relative marginal valuations of health implied in such a system can be indicated by the prices individuals would be willing to pay for additional health and these relative prices are shown by the slope of line such as the one I have labelled *PP* in quadrant I. If these were indeed the implied relative marginal valuations of health in our community of two people, then *E* on the health frontier is the allocative efficiency point and this, as you may readily see, would entail a distribution of the health budget between A and B shown by point *X* on the budget line on quadrant III. (I have used the 'implied' just now because markets will not, of course, directly reveal the marginal values put upon *health*, but those put upon *health care*. The 'shadow price' of health can, however, be inferred from these, given the production functions.)

However, I do not myself much like this way of extending the idea of efficiency, because I think the sort of principle that ought to guide interpersonal allocation ought to be much more guided by judgements of fairness or equity. I would prefer to see the idea of efficiency in health care as being to do with the supply side, and this seems to me to be the way the government also sees the matter. I personally tend to prefer points closer to *e* on the health frontier in quadrant I, which, since it lies on the 45° line through the origin in quadrant I, indicates (complete) equality of health between A and B. Less strongly, I prefer points on the frontier closer to the 45° line than the starting point *S* is. Note, incidentally, that aiming at greater equality of health will generally involve an unequal allocation of resources between individuals (tracing back from point *e* to the budget line does not bring you to the half-way point along it, X_A is not equal to X_B). Note also that going for equal health is not the same as going for equal health *gain* – which would involve preferring a point as close as possible to where a 45° line passing through *S* cuts the health frontier.

Now it is not my purpose to become bogged down in theoretical niceties, but the analysis we have just done does enable us to make some quite important distinctions (for example, about efficiency in production compared to efficiency in interpersonal allocation, or that going for a more equal distribution of *health* in the community may involve quite unequal allocation of health *care* resources within the community, or that the final distribution of health in the community depends upon presenting health states – and the non-health care determinants of these, the relevant production functions, and the distribution of health care expenditures) and it also enables us to frame a discussion of the efficiency or inefficiency of the NHS more carefully. In particular, it becomes clear that, *if health maximisation is the objective,* particular attention needs to be paid to the sorts of technologies that are used in health care and that, especially, the system needs to be designed so far as possible as to maximise the appropriate use of cost-effective technologies (in the broadest sense of 'technologies').

'Macroeconomic efficiency' revisited

What of our authors' concerns about 'macroeconomic efficiency'? This can now be seen to be a question of the location of the budget line in quadrant III – greater resources for health care will push it out further from the origin and fewer would move it towards to origin. As it moves away from or towards the origin, so the health frontier moves away from or towards the origin in quadrant I, *assuming that efficient technologies only are used.* And here is the rub. In a system like the British one, governments will be reluctant, quite apart from any narrow mindedness from the Treasury, to expand the health care budget if the increases went into income increases for existing resource owners (such as doctors and nurses) rather than additional inputs, or if any real increase

in resources increased the use of ineffective health care technologies. In such cases, the budget line might move out but the use of inefficient production functions would still result in an outcome point beneath the health frontier in quadrant I of the diagram – and it may even cause the frontier to shift towards the origin. In these cases, additional health care spending yields no ultimate additional benefit to the clients of the system.

Our authors note that does not seem to be much of a correlation between international health care expenditures and their measure of health status. There are at least three reasons why this may be the case:

(i) some systems (the US notoriously) afford (some) workers in the health sector substantial rents in the sense of higher incomes than a truly competitive market would permit (here one should never lose sight of the happy symmetry between income and expenditure – more expenditure on health care always and invariably means an exactly corresponding increase in incomes for those in that sector, so calls from those within sectors for more *expenditure* may equally be seen as calls for more *income*);

(ii) systems differ greatly in their expenditures on relatively ineffective (including grossly under-utilised) technologies (again the US is notorious);

(iii) *current* health status cannot anyway be expected to be much influenced by current health care expenditures, even if effective, partly because current health is determined mainly by factors other than health care, partly because the beneficial effects of health care emerge only after the passage of time, partly because current health status is mainly determined by factors occurring in the past rather than currently, and partly because the measures of health themselves are rather poor, being both crude and incomplete (I mean no criticism here – the availability of data limit what you can do at the macro level).

Is the treatment appropriate?

I do not propose to take each of the elements of their Prototype in turn. Space forbids that. Let us, however, note the general character of their proposals.

(i) A Guaranteed Health Care Package

The first striking feature is the proposal for a Guaranteed Health Care Package (GHCP). I do not have much quarrel with that but, then, there is not much to quarrel about since any quarrelling would be bound to focus on contentious issues to do with the contents of the package. It may be a good idea for the NHS to adopt a GHCP (through its definition would be no less contentious).

(ii) Compulsory competitive health insurance

The second striking feature is the proposal for competitive insurance coupled with a much more aggressive use of co-payments. The arguments for this seem to be advanced on efficiency grounds, though they are not systematically set out. Where in terms of the diagram are the efficiency gains likely to lie? One possibility would be that our authors expect their proposals to generate a more optimal overall expenditure on health care, which I take to mean an increase! They expect the budget line to move out and that this will, in turn, push out the health frontier. Well, since the compulsory two-part premium is effectively an ear-marked tax, it is certainly possible that this will happen. But it would depend, of course, on political judgements about acceptable premium levels and public subsidy, and we really need a theory of public choice to enable us to say how government would exercise its discretion. In any case, if there really are grounds for optimism in this respect, would it not be simpler to create a National Health Service Fund, hived off from central government, which would receive the earmarked income-related premiums of the public, and which would then allocate this to existing purchasers? That would create as much 'openness' as our authors' suggestions, enable the public to express their views on the adequacy of health care spending at election, and also enable a reduction in, say, income tax yield equal to the current expenditure on the NHS.

They also expect that such a mechanism would enhance the productive efficiency of the system by weeding out inefficient technologies. I cannot myself see the mechanism by which this would happen, unless the new insurers were somehow more effective purchasers for the needs of their clients than current purchasers. But why should

they be? In my analysis, the main reasons for inefficiency of this sort lie in the availability of information on best practice for maximum health gain and an environment in which purchasers have the ability to create incentives for providers to use best practice and providers the means of controlling (mainly) physicians so as to adopt it. Our authors have nothing to suggest that is additional to current policy or structure in this respect.

Or it may be that our authors think that a restructured insurance function would move society to a more desirable point on the health frontier than where they think we currently are. Well, here it needs to be said that premiums are not themselves the cost of using the service and so they themselves will not cause much moving, though they may have important (and I would guess regressive), consequences for the sharing of the burden of health care finances.

I think the Study is frankly naive about the workings of what the authors hope would be competitive insurance. Although they tell us nothing of the billing and monitoring mechanisms to be used, it is quite clear that the transactions costs of competitive insurance are bound to be high (I set aside the costs to government of monitoring and regulating the industry in order to maintain its competitiveness). At least, high in comparison to the single monopolist insurer in the form of the government itself or an agency to which the insurance function has been allocated (I presume through competitive tender!). They minimise the dangers of cream-skimming, which I conjecture would be substantial. It is very easy for a company to turn away potentially unattractive clients (for example by having user-unfriendly application forms, unhelpful responses to telephone enquiries, discourteous front office staff). If I were to run such an agency I have absolutely no doubt that I would easily be able to erect informal mechanisms to cream-skim in ways impervious to any regulatory correction and which, at the end of the day, when my bottom line results raised eyebrows, I would be able to claim, without fear of authoritative contradiction, simply reflected my superior efficiency coupled with the free exercise of consumer choice. But it would all be mostly sham of course! At any rate, detecting my sham would involve a costly, and presumably public, bureaucracy.

(iii) Use of co-payments

Would the more 'rigorous' use of co-payments for 'consumers' (including co-payments for components of the Guaranteed Health Care Package) shift the distribution of health in a way most of us would think desirable? Further, and more fundamentally, who is the consumer? One of the odd things about the economics of health systems is that nearly everyone agrees that the principal character who determines what resources shall be used, and for whom, is a doctor. This is only to be expected. Patients have very little understanding of medical technology and effectiveness (even less than doctors!). They are usually worried and anxious at the time of consultation, most are elderly, many are confused, and many are frightened. While they have the right to have their values and personal circumstances understood and respected by their doctors (this is, after all, one of the main reasons for having a system of GPs) in most cases *decisions* about resource commitment are actually taken by the doctor. Patients may *receive* care and, in that sense, be consumers, but they are unequal partners in the *decision* to consume. In this context it is odd to extrapolate from other walks of consumerist life, when personal judgements about what to buy are much less clouded by fear and anxiety, and are less likely to be delegated in whole or part to a professional agent (who may have his or her own personal agenda to pursue which may conflict with the patient's) and, via this extrapolation, suggest that financial *brakes* (note we are now trying to reduce expenditures, not increase them) be applied to the *patient*. Why not the *doctor*? What are the grounds for supposing that the greater use of *patient out-of-pocket payments* would enhance the efficiency of the system? None are presented and I cannot imagine what they might be. Isn't it all a bit like blaming the overcrowding in prisons on the absence of co-payments for their use? The analogy may not be perfect – after all, people are not sent to prison in the same way of for the same reason as they are sent to hospital. But it is apt in the sense that the decision is largely taken by another party.

Would it not, therefore, be more sensible to charge the doctor? After all, he or she is the real decision taker and he or she is in a much better position to form a judgement about the legitimate claims of one patient relative to another on the inevitably

limited resources available to maximise the community's health. But, then, is not that what we more or less have with fundholding (GPs with budgets to purchase health care at prices set by other providers) and increasingly in the internal budgeting systems in use in hospitals? If it ain't broke, why fix it? I am sure that our authors do not intend this but there is an unmistakable whiff of victim blaming in these proposals. The patient is a victim of ill-fortune or self-induced calamity, so let's saddle him or her with further burdens, even though there is no perceptible reason for supposing that these burdens will, even in subtle indirect ways, lead to substantive improvements in either welfare or health. I do not object to modest charges, mainly as fund raisers. But 'rigorous' charges? There is disturbing evidence, especially from the USA, that even minor user-charges subject to a modest annual maximum annual outlay per insured person, deter – and deter particularly utilisation by children and the poor. If we are to deter people from early consultation with GPs, which is the stage at which preliminary (and sometimes final) judgements about the need for medical care are made, then there is the grave risk of introducing what would actually be a feature that would substantially impede the system's ability to deliver health gain – for it would become increasingly difficult to *identify* the existence of the very needs the system is there to meet, let alone set about meeting them. But, then, our authors do seem to have it in mind to have extensive exemptions from copayments, even in a 'rigorous' system. But how would that differ from what we now have? Whose consumption are they trying to deter? How rigorous is 'rigorous'? And whose bureaucracy would manage a system with exemptions?

A part of the Study's case for co-payments is 'to make patients more aware of the cost of treatment' (p 1133). I find this argument at best incomplete. It is incomplete partly in that it is not clear what would follow in the way of behaviour from such awareness, apart from a normal responsiveness arising out of any elasticity in demand, which is likely, as I have argued, to impair the cost-effective maximisation of health gain. It is incomplete further in that paying only *part* of the cost at the point of use is in fact to receive a *false* signal about (marginal?) cost. We already know that much of the British electorate thinks that it has 'paid its share' of the public expenditure on health care via National Insurance. If anything, then, the message received by patients would be that the care they receive cost much less than it actually does. This proposal does *not*, then, produce the transparency claimed for it.

(iv) Enforceable contracts

The authors' advocacy of enforceability seems much too bold in our current state of knowledge. It is not clear what the relationship between purchasers and providers is in terms of contractual obligations, statutory obligations, and obligations arising out of tort and restitution. Additionally, effective contract enforcement is crucially dependent upon information being available that will stand the test in determining whether or not obligations have been carried out. Such information concerns, among other things, information about service mix, quality, and risk. A further complication is that the status of patients in the contracting process is problematic. The traditional contract doctrines of privity and consideration preclude third parties like patients from enforcing contracts even when such contracts are made to further their interests. Even when all these issues have been resolved there would remain the question of the behavioural and economic implications of them.

It therefore seems clear that much more experience and research is needed here and I frankly doubt whether the general conclusion to which we might eventually come would be to make all contracts legally binding. The current arrangement is that contracts between purchasers and providers are not legally enforceable as *contracts* but are subject to arbitration by the Secretary of State. This is itself a murky legal area and suggests, along with all the other considerations, that there is much to be yet thought about concerning the legal status of contracts.

(v) Insurance, moral hazard, externalities and agency

There is a good deal more that might have to be said once the proposals were got up in greater detail. For example, would premiums be set for individuals or families and whose income would

count in the income-related bit of the premium, whose health experience in the health-related part of the premium? Are these matters which could be left to the market to sort out? If the government chose to have a health policy, as seems reasonable to suppose it might, what would be the mechanisms by which such a policy might be implemented? How would the vexed interface between the health services, conventionally defined, and local authority services be managed under the proposed reforms? We have recently seen a major switch in the location of care away from institutions to the community – 'there's no place like home' (even if it's in a cardboard box under a railway arch). How would such policies be developed and managed under the new system?

All systems of health insurance have their own ways of resolving moral hazard problems of various kinds – those that arise *ex post* at the consumption end when insured parties have an incentive to demand more when the user-price falls as a result of insurance, those that arise (also *ex post*) when providers see opportunities for billing practices that inflate the true costs of effective care, and those that arise *ex ante* when insured parties face a reduced incentive to avoid the circumstance that may lead to their making a claim on the insurer.

There is an undoubted potential inefficiency inherent in moral hazard – 'potential' because although moral hazard tends to increase consumption, whether it does so beyond optimal rates depends on the extent of externalities that lead the social optimum rate to be one higher than the individually selected rate. Although the authors refer to this phenomenon (p 144), they claim that there are few such effects other than that of communicable disease. This sort of eternality has been largely internalised by the NHS and public health measures in the UK, as have the utility interdependencies to which they merely refer in a footnote on the next page. However, the fact (as I conjecture it to be) that these have been largely internalised does not, of course, mean that they cannot be 'uninternalised' if the system were to be changed. These effects are thus a potential (and potent) source of *inefficiency*. The standard market response to moral hazard (of the consumer's *ex post* kind) is co-payment but we have already seen that there are grounds for doubting the relevance of this mechanism in a system aiming to maximise health.

What would be the future role of GPs in the reformed system – still gatekeepers, still the coordinators of community and institutional care, still those professionals chiefly charged with the task of knowing a good deal about the 'whole patient' and able to make clinical (and other) judgements in the context of as a complete an 'agency relationship' as probably exists anywhere in the world? Would they still be those who purchase from the secondary sector for their patients? Or will the new system require the patient to make an initial diagnosis to decide whom to consult (pain in the shoulder therefore I shall go to a physiotherapist), give him or her direct access to outpatient clinics in hospital, and require him or her to make their own arrangements for after care, community services, and all the rest?

And what of the government's own insurance fund – the NHF? (p 1130), which is supposed to operate on a level playing field with the private insurers, and without public subsidy, but which seems likely to wind up with a highly unbalanced portfolio of risk, if only because the poorest will be the sickest?

I do not ask these as rhetorical questions but as ones that need addressing before a full evaluation of the proposals can be made. And they need addressing because they are important matters that have the potential for major disturbance both to the efficiency and the equity of health care in the UK. Nor do I ask them in the fond belief that the NHS as it is, and as it is now evolving, has found the perfect answers. I ask them because the onus is on those who propose change to be clear about what's wrong, how significantly wrong it may be, what's needed to put it right, and how much improvement it may be reasonable to expect as a result. Reformers do not have to promise the earth, but we do need to know the approximate shades of green of the grass on the other side. And the burden of reasonable proof lies with them.

(vi) Patient choice

Another principal objective of the NHS since the recent reforms has been to widen patient choice. Here we confront a large number of difficult issues, most of which have to do with the 'doctor-patient relationship' and the 'agency role' of doctors. In its idealised form, this relationship consists of two

individuals coming together to determine a course of action. The doctor is supposed to bring to this relationship an expertise in the probable consequences of alternative courses of medical action and a skilled judgement as to what procedures are likely to be effective. The patient brings personal circumstances, values and preferences, perhaps occasionally some medical knowledge, and is frequently confused, frightened and having difficulty articulating his or her perceptions, even to a GP with whom they may be very familiar. The art lies either in the patient transmitting the relevant personal circumstance and values for the doctor to fit them into his or her portfolio of medical knowledge so as to make a recommendation, or in the doctor transmitting the medical knowledge for the patient to fit into his or her portfolio of personal knowledge so as to make a decision.

Several things are required if this relationship is to work well:

- the first thing is that the doctor be thoroughly competent in his or her expertise;
- the second is that the relationship be such as to encourage trust on both sides;
- a third is that the resource and reward environment in which the doctor operates should not cause his/her advice to be compromised by factors that are not a legitimate part of the relationship while, on the other hand, enabling him/her to form a judgement about the priority that claims on resources by other patients of their own (they are not typically in a good position to judge the claims made by patients on other doctors relative to those made by their patients on themselves) and which can often involve the tactful denial of care of some types to some patients whose need is judged to be of very low priority);
- a fourth is that the willingness of the patient to come forward to have questions answered, anxieties allayed and needs assessed should not be prejudiced by irrelevancies such as co-payments;
- a fifth, on which I want momentarily to focus, is that the balance between the weights given to the doctor's and the patient's judgements must vary according to circumstances.

There are some cases where the doctor's judgements are extremely marginal, for example, shall the patient have a private room in hospital with bedside office facilities? In such a case, and assuming that having or not having these facilities really is irrelevant for the medical outcome, there is every reason to permit free choice, out of pocket payment, and private supplementary insurance, there being no obvious threat to either health or equity from the exercise of such choice. Such possibilities are clearly implementable within the current public and private arrangements. In other cases, the decision must be balanced between the two. For example, in many situations when the question arises as to whether a woman shall have a caesarean section, and especially in difficult decisions where there is a relatively low risk of a good outcome from a particular treatment and quite a high risk of a bad outcome – the doctor may be quite good at judging the *risk* but the patient is more expert in judging the *acceptability* of the risk. Another case requiring balance is where there are difficult trade-offs, for example in the case of cancer of the larynx, where surgery may prolong life briefly but at the cost of the loss of voice and medical management may involve a shorter life expectation but use of the voice for longer.

At the other extreme are choices where the patient is in no position at all to contribute to the decision, as when he/she is unconscious and an immediate decision is required (though relatives may be legitimately involved in lieu). Some limitations of choice may be more damaging to patient freedom and autonomy than others. It is a characteristic of some systems (for example, Preferred Provider Organisations in the US) that there is a limited choice of hospital doctors from whom to choose. This undoubtedly reduces the range of choice but, if there are good other grounds for limiting choice in this way, the loss may be judged acceptable provided that the controlling doctor (say, the GP) has confidence in those secondary doctors to whom referrals may be made and the patient too has a similar confidence in the GP's judgement. Systems which (or doctors who), however, deny patients the opportunity to have their values properly taken into account are unacceptable. Systems which arbitrarily deny choice (even when there is a willingness to pay) over the quite considerable range of 'hotel' type services which necessarily

accompany much medical care, are likewise unacceptable. The NHS plainly has a long way to go in developing the latter freedoms of choice. But creating these opportunities requires no great radical reform.

Which are the choices which our authors wish to see expanded and why are their proposals needed in order to bring this expansion about? Again, these are not rhetorical questions. They need answering before one can enter in to a useful dialogue on the meaning of patient choice, those elements to which greatest importance is attached, the principal deficiencies of the NHS that need putting right, and the various means at our disposal for putting them right. I miss such a discussion in the Report. Indeed it seems quite likely that the introduction of competitive insurance (and fee-for-service physician remuneration?) could all too easy militate against the ideal relationship, as I have described it, for example, by directly encouraging the use of treatments that are to the doctor's but not the patient's advantage, or indirectly by encouraging the hospital sector to invest in 'me-too' technologies that involve hospital doctors operating too low on their learning curves for effective (let alone cost-effective) care.

The short term proposals

The short term recommendations (p 1124-6) are (I omit any that are the same as long term ones):

- increase the rate of introduction of capitated GP and Health Authority purchasers;
- increase overall funding;
- introduce an earmarked tax called health premiums and reduce other tax equivalently;
- depend less on block contracts and use prospective cost per patient instead;
- increase the use of co-payments to include primary, secondary and tertiary care (with appropriate exemptions.

I tend to support the first three of these and reject, for reasons already rehearsed, the last one. I am not sure what to make of the authors' contracting proposals. On the one hand they want to stop block contracting; on the other they want greater flexibility. No discussion of optimal contracting can ignore the transactions and enforcement costs that are entailed. It is odd that block contracts, which *may* be optimal at least for some packages of care, are to be outlawed. In general I incline to an evolutionary approach: not knowing a *priori* what is right one must rely on experience.

I ought, however, to say in relation to the first recommendation that there is a major tension in the present system between purchasing Health Authorities and purchasing GPs. The former are charged with identifying the needs of their catchment communities (which I regard as the single most important feature of the recent reforms) and arranging for it to be met; the latter deal with the needs of their own patient group within the larger group. It is plain that coordination and consistency are required and that the ability of Authorities to discharge their duties was being increasingly prejudiced as their income was topsliced to fund GPs. It would have been interesting to have read an analysis of this issue in the Study and what role they see in the future for Family Health Service Authorities, whose merger with DHAs is the current policy response to this problem.

Although the Prototype seems to envisage the removal of purchasing District Health Authorities, the specific short term proposals for the UK allow for their retention, with 'consumer' choice of Authority. It seems, then, that we are to envisage three sorts of purchaser: GPs, Health Authorities and competitive insurers. This seems to be a recipe for chaos, especially if the HAs were also competing for clients' custom. How, for example, would differences be reconciled if, as seems likely, GPs and competitive insurers felt obliged by market pressures to purchase services judged ineffective by Authorities, or that did not address the major needs identified by Authorities? And what sort of information base is to be supplied (and by whom) to inform better the purchasing choices of GPs and competitive insurers, granted the already inadequate base that exists for purchasing Authorities, the very wide variability of GP competence in epidemiological understanding, and the general ignorance of the public for whose custom these various agents will be competing?

There is something of a *curiosum* in the proposals to which I have not referred so far: that there

should be price ceilings on some forms of treatment, such as geriatric care (p 1133). This seems odd, in part because geriatric services (or at least the health care services used by the elderly) are such a large part of total expenditures and the proposal seems almost an after-thought, and in part because it is so out of line with the market orientation of the rest of the Study. What shall these ceilings be, who shall set them, and who monitor and enforce them? What, anyway, is their justification in terms of the economic efficiency that underlies the whole Study?

Final comments and conclusions

I have focussed my comments mainly on the 'long term' treatment recommended for the patient by the authors rather than the intermediate treatment they recommend. The reason for this is plain – should the long term treatment be deemed not appropriate or cost-effective, then we need not enquire too diligently into the intermediate treatments. Moreover, the short and medium term proposals are in large part contained within the long term ones, so any discussion of the latter will have embraced a good deal of the former. I must say, however, that I have grave reservation both about the diagnosis and the recommended treatment. My analysis has not been made the easier for the absence of a clearly defined set of objectives. I have tried to be rather clearer and specific in my own reactions. My own view is that the patient's condition is not such as to warrant the draconian measures proposed and that even if it were in a parlous state, I am not convinced of the efficacy or cost-effectiveness of the treatments recommended. Nor am I convinced that the objectives sought, once they were made clear, are those to which most of us would want to subscribe.

There are also many gaps in the analysis. Changes as radical as those proposed for the UK generate a host of important questions – to some of which I have alluded – and which would require a good deal of careful thought and investigation if the dangers inherent in them were not to swamp potential benefits from such changes, if there are any to be had. Issues such as the definition of 'need', how best to promote effective medical care, the transaction costs of competitive insurance, the sorts of patient choice that need expanding, adverse selection, moral hazard, cream skimming, externalities, equitable distribution of benefits, enforceability of contracts, and so on, are complex and best discussed in the context of concrete proposals and an awareness of the nature of the problems that can arise.

Nor, I fear, ought anyone to take the limited forecasting exercise presented in the Study (p 73-80) too seriously. A good idea of the predictive power of estimating equations can be given if a subset of the data are used to estimate the equation, whose predictions for other years in the data set can be compared with the actuals for those years. It is not entirely clear what our authors did from the Report, for the estimation period of the 'supply equation' is stated (under the equation) to be 1960-1986, but the text refers to an estimation over the period 1960-1990. There is a comment that the equation for the shorter period does not explain the period 1986-1990 well, and this is not surprising since these equations have been shown to be unstable elsewhere. Their estimated income elasticities are higher than most of those found in the literature, which gives grounds for caution. There has been a considerable discussion of these procedures in issues of *Journal of Health Economics*, which is not referred to, where methodology and the literature are discussed in greater detail. Moreover, the projections of 'need' are based *not on UK data but on those for other countries*, on the basis of which it is said (p 78) that UK current need is for 9 per cent of GDP to be spent on health care rather than the current 6.1 per cent (1990). I have commented earlier on the weakness of this kind of comparison but to use other countries' estimates of need as a proxy for the UK's seems extraordinary. Taken together, these considerations suggest that the somewhat alarmist warning (p 80) about an increasing and unsustainable shortfall between need and supply ought not to be taken too seriously. If there is a shortfall it needs to be identified and detailed in other ways – ways which have a closer relationship to the economic concept of efficiency and the ethical concept of equity.

My own view, to put it rather generally, is that current policy towards the NHS is, in broad terms, right in terms of *structure* – what I have called 'demand-side socialism', which is, of course, quite consistent with private ownership (and for-profit

motivation) on the supply side, provided that providers act at the behest of purchasing Authorities and GPs. I think the objective of maximising health gain is appropriate and that the separation of provider and purchaser has had, and is having, a useful effect in forcing attention (especially purchasers' attention) on issues of effectiveness and need and is galvanising the research and development community into the provision of a demand-led set of methodologies and results that, over time, stands a good chance of revolutionising the overall efficiency of the system. I am less persuaded that competition between providers is likely to yield efficiency gains of significance apart from the sort of relocation of activity from high-rent sites to low-rent sites that is currently raising difficult questions of 'exit' for policy makers. I also take the view, however, that the many distributional issues in health and health care (both on the financing side and the delivery side) need much greater thought, discussion and research, particularly at the policy-making centre and at the level of the purchaser.

I do not think that the case for radical change in financing methods has been made, though I do feel sympathy for the idea of hiving off the compulsory public insurance function. An independent National Health Fund funded out of earmarked contributions proportionate to average income tax rates has attractions, particularly if it could be so designed as to cause it to fall outside public expenditure. I, however, would not see this as a residual sort of fund of the kind of envisaged by the authors. There must be some way of operating a compulsory insurance system, monitored and regulated by central government, that was not a part of the tax system (even if the Inland Revenue was the agency via which – for a fee – the contributions of taxpayers were collected and delivered to the Fund). This would not completely insulate health care expenditures from the probably arbitrary limits imposed by macro economic judgements, and it would evidently fall to government to determine or control the premiums charged, but it would – or it could – create that magic transformation through which ('unproductive') public spending was turned into ('productive') private spending, and 'non-wealth-creating' production was transformed into 'wealth-creating' production. This is, of course, merely to swap myths. Unfortunately, myths are hard to get rid of but they do have consequences, some of which are bad and some good. Replacing a myth with bad consequences with one that had good consequences would be no bad thing and the main good consequence would be to give the public an opportunity more directly to determine the total spend on health care.

I am therefore much more modest in my own proposals for reform. I prefer the UK Cochrane Centre to competitive insurance as a means of securing greater efficiency, public purchasing health authorities to private health insurers as a means of revealing need, and GP gatekeepers to co-payment-determined independent access to the entire system as a means of investigating *prima facie* need and coordinating the work of providers for individual patients. I think health care is probably under-financed in the UK and would welcome a new mechanism to correct this provided that increased financing was translated into increased cost-effective resource use and not rent-seeking and waste. In short, I think what we have in Britain is a nut of a problem and our authors are offering us a sledgehammer which not only smashes the nut to smithereens but may also break our toes, or backs, or both. It also seems altogether premature to jettison the current structures when they have hardly yet had time to deliver on their promise. The prospect of subjecting the NHS to yet another upheaval is too awful to contemplate, even thought the evidence is that managers would respond as energetically as they have to *Working for Patients* and all its *sequelae*. (In passing, one might note that the alleged resistance of the NHS 'bureaucracy' to change is quite unsustainable. One can only marvel at the way in which NHS management has responded to the recent challenges of managing change).

Let me end with some wise words from C E Lindblom: ***'A market is like a tool: designed to do certain jobs but unsuited for others. Not wholly familiar with what it can do, people often leave it lying in the drawer when they could use it. But then, they also use it when they should not, like an amateur craftsman who carelessly uses his chisel as a screwdriver.'***