Paying for health

Sir Douglas Black  University of Manchester

Author’s abstract
Health care systems, irrespective of how they are financed, present the paradox that to some observers they appear as a major component of social benefits, while to other observers they seem both excessively costly and limited in their effectiveness. These differing perceptions may be explained in part by the diversity of the determinants of health and disease, only some of which are amenable to those preventive or therapeutic measures encompassed in a health care system – the majority of determinants being genetic, societal, or else uninfluenced by those interventions at present available within a health service.

The share of national resources which should be devoted to health care, and the method of raising resources, are primarily matters for political decision; but a national system has advantages both of economy and of comprehensiveness. But when it comes to allocation of resources within the established health budget, the knowledge and skills of health professionals are essential to informed decision-making. The possibilities depend critically on the ‘state of the art’ at a given time, as is illustrated by the radical changes over time in what could be done for patients with renal failure; and health professionals are likely to be most aware of current options, and of how to choose between them. More speculatively, they are also less likely to confuse the attitudes appropriate to providing a service with those required to run a business.

In many countries, the cost of health is a matter of increasing concern, not simply to health professionals and managers, but also to politicians and the media. Advances in medical knowledge and skills, whether they come from stepwise increments in knowledge, or from the uncommon ‘breakthroughs’ so dear to the media, are making possible realistic prevention or treatment for a widening range of diseases. For these particular diseases, modern medicine is highly effective; but it is also more sophisticated, and thus more expensive, generating costs which have to be set against the economic benefits which flow from successful treatment. And there are still very many diseases, notably tumours, arthropathies, and arterial diseases which are often not amenable to radical treatment, and which cause high mortality and even greater morbidity. The very success of medicine in younger age-groups adds to the burden imposed on services by the more vulnerable elderly; and in many countries this accentuates an underlying demographic trend towards an increased proportion of elderly people in the population. Concern over the economic consequences of these developments has enlisted the interest of professional economists, and there is now a flourishing discipline of health economics. It is important to distinguish the contribution of a health economist from those of an accountant, of a manager, and of a health professional. An accountant is expert in financial costs, but not in the assessment of benefits, particularly those which are difficult to translate into money values, such as increased perception of well-being or release from anxiety. A manager is concerned with process, whether it be of systems or of staff; and not with outputs which he is ill-qualified to assess, but which are nevertheless of prime concern to the patients for whom the service is designed. A health professional practises the art of the possible within the scope and limitations of his expertise. But the health economist has to take a view which is in some ways broader than any of these, taking in not just the costs of process, but also the values of outcomes. And he has in some sense to be a trustee for tomorrow’s patients, in face of the seemingly inexhaustible demands made by patients of today, often most stridently not by patients themselves but by their self-appointed proxies. His is not of course a self-standing, free-range discipline – he depends on the accountant for an assessment of costs; on the manager for maintaining a system which will provide the information needed for economic analysis; and on the health professional for an analysis of available options and for categorisation of disease states and their likely outcomes. As someone who is primarily a physician, I am conscious of my rashness in entering the domain of the health economist; but I have always looked on health economists as natural allies of clinical doctors and nurses, so I look on my incursion as one into friendly territory.

And of course in this paper I can cover only a small part of that territory. I limit myself to three main themes – the determinants of health, some of which

Key words
Health systems; health economics.
rather happily turn out to be ‘free goods’; the share of national resources which it is reasonable to devote to ‘health’, and the criteria for making choices within the (necessarily limited) resources available for health.

Determinants of individual health

The first down-payment on our health is made some nine months before we first see the light, and whatever element of choice there may be in it, it is certainly not our own. That payment, vicariously made, gives us a quantum of health capital which is a major determinant of the health experience which we may subsequently enjoy. Our health capital at fertilisation may already be reduced by the seeds of future specific ‘heritable diseases’, or by a multifactorial genetic endowment which threatens our life-span through a variety of handicaps such as an increased liability to raised blood pressure or to a form of diabetes. Nor of course is the womb the safe haven of popular belief – many a pregnancy ends spontaneously in abortion or stillbirth, not always due to intrinsic non-viability of the fetus, but sometimes occurring in women who are to all appearances healthy as well as in women with specific disorders such as renal failure. Some congenital malformations and some forms of mental subnormality can be related to prenatal influences which pertain to the intra-uterine environment rather than to genetic endowment (1). Particularly before we are born, the distribution of responsibility for good or bad health between the seed and the soil is often a matter of speculation; and also of little pragmatic consequence, beyond the provision of the best antenatal care and appropriate clinical genetic advice, which are of course important in their own right.

Our state of dependency persists after birth into infancy, and in diminishing degree into childhood and adolescence. The environment in which we grow up, and which to a large extent now shapes our lives, is partly natural (and no species has a greater range of habitat), and partly societal (parents, siblings and peer groups, schooling and social class). Accidents, infection and sadly in many parts of the world undernutrition and malnutrition, are the main threats to physical development. Mental development requires appropriate stimulation as well as native intellectual capacity. Many of the things which influence our physical and mental development are beyond our control; but as we become mature, or at least more mature, we begin to have perceptions of autonomy. In the matter of health, we appear to have choices in such matters as what we eat, whether we smoke or drink, and what risks we customarily accept. The choice may not be as autonomous as it is perceived to be by the chooser – there must be an element of social determinism (family circumstances, peer group influence), perhaps even inbuilt cravings for food or mood-changing agents. But unless we are completely deterministic in our outlook, we must, after all allowance has been made for what is beyond our control, accept that we are capable of making at least some of the relevant choices, and by the same token accept the moral responsibility of doing so.

The main theme of this paper relates to populations or societies, and not directly to individuals. Why then should I choose to preface it with an outline picture of the influences on individual health? Most obviously, because any population is an aggregation of individuals. But also, to give at the outset an indication of the complexity of the issues involved, so as to give a ground for deprecating simple solutions; and to affirm that societal provision, important though it is, is not enough – we also share in the determination of our own health.

Determinants of health in populations

From the time of Hippocrates, it has been known that health is affected by ‘airs, waters and places’, influences which will potentially affect all those exposed to them. However, the effect of such general agents on particular people will be greatly modified by individual susceptibility, and this too has been known at least from the time of Pythagoras – not everyone who eats broad beans will suffer from favism. There is thus a dynamic interaction between factors to which the whole population is exposed, and important characteristics of the individual, which may be constitutional (for example, compromised immune system), behavioural (so-called ‘life-style’), or societal (social/occupational status).

The reality of life in primitive societies is probably closer to the picture of Thomas Hobbes ("the life of man, solitary, poor, nasty, brutish and short"), than to Rousseau's 'noble savage'. Sadly enough, the scourges of war, pestilence and famine are still with us in many parts of the world; but for our present purpose, an historical perspective seems more appropriate than a geographical one. There has been registration of deaths in England and Wales since 1838; and before that there are records of population from which changes in mortality can be inferred. Thomas McKeown (1) has made these records the basis of an influential analysis of the factors which may have contributed to the steady increase in life-expectancy which has taken place from around 1700 onwards. For the earlier period before 'causes of death' were attributed, he ascribes the improvement mainly to improved living conditions, and a better diet, and also – surprisingly – to decline in the practice of infanticide in poor families (Disraeli is said to have believed that infanticide 'was hardly less prevalent in England than on the banks of the Ganges'). The decline in mortality is better documented over the past century and a half – the death-rate fell from 23.3 (men) and 20.7 (women) per thousand in 1871–5 down to 12.0 (men) and 11.4 (women) in 1983 (2). As with the earlier period, McKeown ascribes the greatest effect to improved standards of living, particularly better nutrition giving increased resistance to infection; and also to improvements in sanitation and hygiene, which would also help to reduce infective illness. By comparison
with these social and environmental changes, McKeown attaches less importance to advances in clinical medicine, dramatic though these have been; he points out, in particular, that these have only become significant long after the decline in mortality has become well established. He does not deny the importance of a high standard of medical practice in relieving suffering, and in ‘curing’ acute episodes of illness, especially those due to infection. But at the population level, he sees the social factors as paramount. When we were studying the great and growing differences in ‘health’ between ‘non-manual’ and ‘manual’ workers (3), my colleagues and I came to a similar conclusion, that any disparity in access to medical care was not so important in causing disadvantage to manual workers as was a constellation of factors associated with social deprivation – such things as poor and crowded housing, which favour infection; poor nutrition, the diet being unbalanced rather than insufficient; greater risk of accident in the home, at work, or in travel; and an unwise life-style, with smoking one important variable between the ‘classes’. Important as these social factors are, they are not the whole story. In the past decade, since McKeown wrote, it has become apparent that there has been a greater fall in mortality in those diseases for which there is effective prevention or treatment than in those diseases for which such measures are not yet available. No great surprise there, but what has become clear is that this effect is of sufficient magnitude to show up in statistics of mortality, and not simply as anecdotal evidence (4).

Let me conclude this section on the determinants of health by giving the good news which I promised in the introduction, that some important determinants of health are ‘free goods’, in the sense that they carry no ‘accountable’ costs. For the individual, there is – for weal or woe – his genetic endowment. And as a member of his group, he enjoys the sun, when it shines; and the air which he breathes – for even in the UK, fiscal enterprise has gone no further than metering the water he drinks or bathes in. There is also a health bonus to be enjoyed by those who are congenially employed in a free society.

**Determination of the global health budget**

Health care in a modern society costs much money, and employs many people. Moreover, its claims on national resources have to be judged in comparison with those of housing, education, and social security, which in themselves can make important contributions to peoples’ health. There is an ‘open-endedness’ about possible claims for resources for health, which is perhaps more easily recognised by those who have to meet them than by those who make them. Respect must go to politicians who have the candour to say that not all claims can be met, particularly if they say so while in office, and not in the comparative freedom of being in opposition. Possibly to the disadvantage of his own political future, Dr David Owen said in 1976: ‘All the evidence there is, both national and international, suggests that if need is not infinite, it is certainly so large relative to the resources which society is able to provide now and in the foreseeable future that we can never hope to meet it completely’ (5).

But recognition of the problem, though necessary, is only the first step towards trying to solve it. Attempts to do so are of three general types:

1) The ‘free market’ or ‘laissez-faire’ approach, in which people make their own arrangements with ‘providers’ for health care. This makes access to health care part of the ‘reward system’ of society, by virtue of which those who ‘do well’ financially are privileged to enjoy more expensive health care.

2) An insurance-based system, in which people in good health make a regular contribution, sometimes supplemented by their employer, towards future costs of health care. Examples are the Lloyd George health insurance scheme in Britain, and ‘Blue Cross’ in the USA.

3) A ‘national’ system, in which resources are derived from general taxation, and entitlement is universal in the sense that every citizen can obtain primary care from a ‘general practitioner’ or ‘family doctor’.

Although directly employed by a university, I did my clinical work in the National Health Service (NHS) for
many years, and had previously practised in the insurance-based system which ran from 1911 to 1948. Before 1948, the State benefit was limited to employed persons, and excluded their wives and children, as well as the largely self-employed ‘middle classes’. Having witnessed the partial removal of the financial hardship caused by illness, I have an admitted bias towards the national system.

It should be noted that these three systems can co-exist in the same country, and generally do. For example, in Britain although the NHS predominates, there is a minor insurance component in its finance; and there is also an appreciable private sector. Conversely, in the USA, with its predominantly private system, partly insurance-based, there is much state-financed provision of health care for ‘veterans’ and the ‘indigent’.

Controversy on the relative merits of a free-market system, based on ‘enterprise’ and ‘liberty’, and a state system based on ‘fairness’ and ‘equality’, sometimes lacks the grace of ‘fraternity’, the third ideal of revolutionary France. A free-market system is blamed for its lack of comprehensiveness, a state system for ‘bureaucracy and waste’. The accusation of ‘waste’ often made in the USA but echoed in Britain itself is at least susceptible of objective analysis. Per capita health costs in the USA are three times those in Britain, and the administrative component of these costs is comparably higher; there may also be a component of unnecessary procedures, carried out for fear of litigation, or simply for gain.

Since there is a trend for more affluent countries to spend more on health, a more realistic comparison of health expenditure should perhaps be made with countries in Europe with comparable ‘gross domestic product’ (GDP) to that in Britain. Table I shows the proportion of GDP devoted to health in seven European countries, including Britain. The comparison does not suggest conspicuous waste in the British system.

If the bold question were to be asked, not why we spend so much on health, but rather why in comparison with other comparable countries do we spend so little, I can suggest a number of reasons. The most important of these is our success or good fortune in having preserved a nation-wide system of primary care, based on universal right of access to general practitioners, who are increasingly being trained to a high standard, and who commonly know both patients and their families. They are thus in a position to modify public expectations of what medicine can (or cannot) do; and to control access to expensive, and it may be unnecessary diagnostic and therapeutic procedures in the hospital sector. More directly economic factors include the relative efficiency of using an established system of raising revenue by general taxation, in comparison with a network of insurance or similar agencies; and the opportunity given by a national system to governments of exercising fiscal control. This opportunity has not been neglected, and must have contributed to our relatively low place in the league of health expenditure; more specifically, it has allowed successive governments, as monopoly employers, to protect health service workers from the moral dangers of excessive affluence.

I believe, in short, that we have a system of health care which is not only economical, but also effective in terms both of general availability, and of contributing to health statistics which compare well with those of countries which spend substantially more on health care. As professionals, doctors play their part by avoiding waste in what they do; but the major decisions on the global health budget are inescapably political. Once the size of the national cake, and the proportion of it which can be devoted to health, have been decided, doctors can and must play a larger role in the choices within the health budget. Some elements of decisions on allocations within the health budget

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>European countries with comparable <strong>per capita</strong> GDP</td>
</tr>
<tr>
<td>Percentage of GDP devoted to health care (1986 figures)</td>
</tr>
<tr>
<td><strong>Percentage</strong></td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td>Netherlands</td>
</tr>
<tr>
<td>Austria</td>
</tr>
<tr>
<td>Italy</td>
</tr>
<tr>
<td>Belgium</td>
</tr>
<tr>
<td>Finland</td>
</tr>
<tr>
<td>Britain</td>
</tr>
</tbody>
</table>

**Note** The comparable percentage for the USA is 11.1, for Canada 8.5, and for Greece 3.9.
remain preponderantly political in a broad sense, for example the proportion of resources to be given to acute services, and the balance between care in institutions and in the community; but even in these medical advice should form a part of the decision-making process.

**Allocation within the global health budget**

When it comes to choices within a fixed budget, a guiding concept is that of ‘opportunity cost’. Simply stated, this says that if you spend a pound in one way, you cannot then spend it in another, thus foregoing the second, perhaps equally or more desirable, ‘opportunity’, and thus incurring whatever ‘cost’, in the shape of loss of benefit, failure to take that opportunity entails. This may seem simple, even simplistic; but in applying the notion to health services, Alan Williams (6) exposes difficulties:

‘The golden rule is that only when we are satisfied that the most valuable thing that we are not doing, is less valuable than the least valuable thing that we are doing, can we be sure that we are being efficient in the pursuit of welfare. I guess we have a long way to go yet.’

That too may seem simple; but embedded in it is the need for a judgement of the relative value of different options. As already mentioned, the size of the global health budget, and even certain major allocations within it, are matters primarily for political judgement, informed where appropriate by medical advice, but not dictated by it. But for those decisions which are substantially influenced by what is currently possible in clinical terms – the ‘state of the clinical art’ – I would contend that professional judgement becomes inescapable, and even paramount over managerial aspirations and the clamour of pressure groups. By saying ‘professional’, I do not mean ‘just doctors’, but the array of professions concerned with providing services, and also economists with an interest in health matters.

How do I justify this emphasis on the professional and economic, in contrast to the managerial and actuarial input to decision-making in situations where there is a material clinical component? I take my stand on the great and increasing complexity of the medical database, which requires informed interpretation; on the need to assess outcomes as well as process, both of which require a professional input; but most of all, on the nature of the service which constitutes the main, if not the sole, justification of any health care system. For illustration of these matters, I would like to turn to the care of patients with renal failure, for which at one time I shared responsibility with colleagues who taught me so much.

**IMPORTANCE OF THE ‘STATE OF THE ART’**

My illustration of the importance of the medical database is in a sense ‘historical’, showing how the options, for treating end-stage renal failure have changed over the years, and with them the economic dimension of what is possible. For convenience, I have set out in Table II the availability and cost of successive options. (Renal failure has many causes, some of which are reversible; but when these have been ruled out, and the

<table>
<thead>
<tr>
<th>Option</th>
<th>Availability date</th>
<th>Approximate cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transplant (identical twins)</td>
<td>1958</td>
<td>(£6,000)</td>
</tr>
<tr>
<td>Intermittent haemodialysis hospital</td>
<td>1960</td>
<td>£14,500</td>
</tr>
<tr>
<td>Intermittent haemodialysis home</td>
<td>1965</td>
<td>£9,500</td>
</tr>
<tr>
<td>Low-protein high-calorie diet</td>
<td>1963</td>
<td>(Palliative measure)</td>
</tr>
<tr>
<td>Cadaver transplant</td>
<td>1965</td>
<td>(£6,000) (£1,500)</td>
</tr>
<tr>
<td>CAPD</td>
<td>1980</td>
<td>£8,000</td>
</tr>
</tbody>
</table>

**Notes** ‘Availability date’ is an estimate of the date at which the option became generally available in the UK. ‘Cost’ is annual maintenance cost, except for the ‘one-off’ cost of transplant, which is shown in parenthesis. CAPD is Continuous Ambulatory Peritoneal Dialysis. Before 1950 there was no effective way of avoiding death, given irreversible end-stage renal failure.
glomerular filtration rate has fallen from its normal level of about 120 ml/min to 5 ml/min or less, end-stage renal failure is reached; and before the ‘fifties’ death in uraemic coma was inevitable. Consideration of costs in isolation would show that a renal transplant, with manageable initial cost, and ‘low’ maintenance cost for immunosuppressive agents, is the ‘best buy’ – and in general it gives better ‘quality of life’ than a regime of repeated dialysis. But it is limited in application by the availability of suitable kidneys, and also by the need not just for surgeons trained in transplantation but also for competent dialysis before and after operation, and during episodes of rejection. Again, a crude financial comparison of ‘hospital’ and ‘home’ dialysis, much to the advantage of the latter, neglects two important considerations. Patients who need hospital dialysis are commonly the most severely affected, with complications which in themselves require expensive treatment; and there is the considerable ‘social cost’ to patient and family of repeated dialyses at home. It is easy for politicians, goaded on by pressure groups, to say – as indeed they did say in the USA – that every patient in end-stage renal failure has an entitlement to treatment. But the confident statement of such an objective is not invariably followed by provision of the means needed to achieve it. The extent to which it can be achieved clearly involves professional judgement and dedication, as well as managerial skills and political will.

**IMPORTANCE OF OUTCOME**

A favourable outcome has two main components – **survival**, which is necessary, but not all-sufficient; and **well-being**, or ‘quality of life’, which is of course much harder to assess, but nevertheless of great importance. One possible, relatively ‘hard’, measure of outcome is the capacity for employment; and some years I applied this in an attempt to assess the ‘value’ of different methods of treating renal failure (7). From past experience with the various methods, the likelihood or ‘probability’ (p) of different states of fitness can be ascertained; more arbitrarily or subjectively, an estimate of the worthwhileness or ‘utility’ (u) of the various degrees of fitness for work can be made. For each method of treatment, the ‘value’ is estimated by summing the products of ‘p’ and ‘u’ for the various possible outcomes. Table III shows some results obtained in this way, and once again the assessment is incomplete – it looks as if there is little to choose between home dialysis and transplant – but that misses out both the burden to the family of home dialysis, and the greater well-being, on average, of those who have had a successful transplant.

A more general attempt to combine survival and well-being in a single measure of outcome is represented by the QALY approach – QALY being an

---

**TABLE III**

**Terminal renal failure**

**Fitness for work after different methods of treatment**

In this table the horizontal rows give, for each method of treatment, the **probability** (p) of the outcome specified at the top of the corresponding column. The final row gives an estimate of the **utility** (u) of each of the outcomes. The final column (Σ pu) gives an overall estimate of the relative ‘value’ of the method described at the start of the corresponding row. In the heading, FT = full-time, PT = part-time, Unemp = fit for some work, but unemployed.

<table>
<thead>
<tr>
<th>Degree of employment</th>
<th>FT</th>
<th>PT</th>
<th>Unemp</th>
<th>Unfit</th>
<th>Σ pu</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method of treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dialysis (hospital)</td>
<td>0.34</td>
<td>0.25</td>
<td>0.21</td>
<td>0.20</td>
<td>0.54</td>
</tr>
<tr>
<td>(home)</td>
<td>0.64</td>
<td>0.16</td>
<td>0.14</td>
<td>0.06</td>
<td>0.74</td>
</tr>
<tr>
<td>Transplant (live-donor)</td>
<td>0.75</td>
<td>0.10</td>
<td>0.10</td>
<td>0.05</td>
<td>0.77</td>
</tr>
<tr>
<td>(Cadaver)</td>
<td>0.66</td>
<td>0.12</td>
<td>0.13</td>
<td>0.09</td>
<td>0.71</td>
</tr>
<tr>
<td>Estimated utility</td>
<td>0.90</td>
<td>0.60</td>
<td>0.30</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>
acronym for 'Quality Adjusted Life Years' (8). The QALY of a procedure (which might be an operation, a course of medication, or even a change of diet) is an estimate of the number of years which it should add to the life of a patient, adjusted by an estimate of the quality of the additional life to be enjoyed. This is certainly an important correction – for example, the frequency and duration of survival of patients who have sustained a stroke may both be increased by treatment, but the quality of the life so prolonged may be low. This is allowed for in the QALY matrix – for example, a patient who could not get out of bed and was also suffering severe pain or other distress would contribute a negative value to the assessment of the procedure under investigation, and likewise a patient who was irreversibly unconscious. Although the QALY approach has the great merit of taking account of 'state of health' and not just 'mere life', there are some drawbacks. The estimate of additional years of life must be the mean of variable individual experiences, within each of which it cannot be easy to distinguish between survival which is part of the natural history of the disease and survival which is directly due to the procedure. This estimate of survival is then 'adjusted' by a subjective estimate of 'quality of life', drawn from a quite small panel of health workers and convalescent patients. Since older patients have inherently reduced expectancy of life, there may be bias against them; but it may be noted that this does not prevent hip replacement from outdistancing both renal transplants and coronary bypass surgery in the 'QALY stakes'. It may not be surprising that QALYs are popular with paediatricians and orthopaedic surgeons, less esteemed by nephrologists and geriatricians. It may also be that the simplicity of a single index may be deceptive; and that the complexity of 'quality of life' may necessitate a multi-factorial assessment, such as that provided, for example, by the Nottingham Health Profile (9). However, if agreement could be reached on the weighting to be accorded to each of the thirteen components of that profile, that too is reducible to a single index – but the 'if' in that clause is a large one.

**NATURE OF HEALTH CARE**

My strongest ground of concern for what I see as an unjustified encroachment of managers and accountants on areas which are more properly and suitably the responsibility of professionals trained and experienced in the provision of health care, whether for individuals or for communities, lies in the nature of the service to be provided. Procedures and attitudes which may be appropriate for managing a chain-store are unlikely to be appropriate in a health service. The 'hard sell' can of course have applied to it a veneer of 'public relations'; but that still leaves it far removed from what is needed for health care, which is deeply personal, sometimes painfully so; and which demands a conjunction of human sympathy and professional skill which cannot be replaced either by a purely cash transaction, or by diktat from a manager (or politician) on high.

Let me summarise in a few words what I have been trying to say. The determinants of health are many and complex, even if some of them are economically 'free'. The large economic decisions – how much for 'welfare', and within that how much for 'health' – are inescapably political, even if the expenditure on 'health' is hidden in the web of private care. But when it comes to the assessment of priorities within the health budget, present trends to transfer decisions from health professionals to managers and accountants are not only harmful to patients, but also massively wasteful. This can be simply shown by looking at health costs in the USA, where commerce rules.

_Sir Douglas Black MD FRCP is Emeritus Professor of Medicine, University of Manchester and a Past President of the Royal College of Physicians and of the British Medical Association._

**References**