THE ECONOMIC AND SOCIAL RESEARCH INSTITUTE

AFFORDABILITY VERSUS QUALITY, EFFECTIVENESS AND EQUITY IN HEALTH CARE: IS THERE A TRADE-OFF?

Brian Nolan

Paper for OECD Conference on
"Reforming Health Systems"

November 1994

Working Paper No. 55

Affordability Versus Quality, Effectiveness and Equity: Is There a Trade-Off?

1. Introduction

"That any sane nation, having observed that you could provide for the supply of bread by giving bakers a pecuniary interest in baking for you, should go on to give a surgeon a pecuniary interest in cutting off your leg, is enough to make one despair of political humanity."

George Bernard Shaw made the point as long ago as 1911, but analysts and policy-makers are still struggling to come to terms with the implications of the fact that health care is not like other commodities. To understand why "affordability" of health expenditure is now such an issue throughout the OECD, and why quality, effectiveness and equity in health care are so difficult to define and measure and yet so important, one returns repeatedly to that insight.

The perceived need to contain the level of health expenditure has been the main driving force behind health sector reforms in many OECD countries over the past decade. Why is devoting an increasing share of national resources to health care seen as such a cause for concern? As well as containing expenditure, most countries - as the OECD Health Reform project has shown - have been trying to develop policies to improve the quality and effectiveness of health care, as well as ensuring equitable access to that care. It has become increasingly clear that policies to attain these goals cannot be compartmentalized: policies and goals interact and may conflict. This paper discusses what we have learnt about these interactions, in particular about whether containing health expenditure has to be at the expense of quality, effectiveness and equity in health care. What are the trade-offs?

We begin by looking in Section 2 at why concern about controlling health expenditure is so widespread, at what people mean when they say that health spending is "too high" or growing too rapidly. The notion of the "affordability" of a health care system needs to be scrutinised at the outset. Teasing out what this might mean leads us to highlight how much perceptions about "value for money", rather than simply aggregates such as the share of national income going on health care, affect views about affordability.

We are thus naturally brought to consideration of the relationship between affordability and the objectives of improving the quality, effectiveness and equity of the health care system. This necessarily involves first discussing, in Section 3, how the outcomes, efficacy and cost-effectiveness of health care interventions can be measured. We then focus in

Section 4 on what can be learned from recent research and experience about methods of improving efficacy and cost-effectiveness, in particular by wider diffusion of the results of that research.

Both efficiency and equity call for a pooling of the risks associated with ill-health. In OECD countries the risks to individuals of incurring substantial expenditure on health care are generally pooled by "insurance" in one form or another - usually bearing little resemblance to actuarially-fair insurance. How should the costs of health care then be spread over the insured population? Since insurance is so central to health care, Sections 5 and 6 focus on the relationships between insurance and quality, cost-effectiveness and equity. This includes discussion of insurance and medical technology, risk-rating, the problems of adverse selection and "cream-skimming", and the impact of competition (within the private and the public sector or between the two) on the quality of care. The central question for policy is how to protect the quality of care, and equity in access to that care, while promoting cost-effectiveness. The objective here is not to provide an exhaustive treatment of this much-travelled ground, but rather to highlight the key inter-relationships and possible trade-offs between goals which are the focus of this paper.

Some concluding comments are given in Section 7.

2. Affordability

The need to contain the overall level of health expenditure has dominated health policy in many OECD countries over the past decade. This has not been confined to countries at the top of the spending league in terms of share of GNP/GDP going on health, though it has been particularly pronounced in the biggest spender, the USA. Countries currently devoting much less of their resources to health than the USA have introduced policy measures aimed at restraining its growth. Often the goal appears to be to bring increases in health spending in line with GNP growth and thus stabilise the share of national income going on health, though in some cases a reduction in that share appears to be the aim. The trends in health spending giving rise to this policy focus have been extensively documented elsewhere, most comprehensively in recent OECD publications.[1] Here our interest is first in why devoting an increasing share of national resources to health care is seen as such a problem - why are the rates of growth in health spending in many countries seen as "unsustainable"? As Newhouse (1992) points out, neither citizens nor economists are especially concerned about rapid growth in most sectors of the economy, like the personal computer industry, the fax machine industry, or the cellular phone industry. (Indeed, the emphasis tends most often to be on the

[1] See OECD 1993, Hurst 1992.

jobs provided by such growing sectors). Why then is reaction to growth in the health care sector so different?

One reason might be that in many countries much of health spending is public spending, and so general concerns about the state of the public finances and the need to cut back budget deficits spill over into the health area. Just as the level of for example social security spending and trends in that spending have been seen as contributing to such fiscal problems, so has public expenditure on health. Resistance to increases in public spending in turn reflect what is perceived as unwillingness to pay higher taxes to reduce deficits or pay for more public services, or concerns about the detrimental effects such higher taxes may have on incentives and economic growth. Particularly from an equity perspective, the socio-political factors involved in determining the level of public spending which is deemed "affordable" at a particular time require careful study

Concern about the level of public spending, however important, does not convince as a complete explanation for the generalised antipathy towards continued rapid growth in health spending across the OECD. This antipathy is shared by countries where the private element in health spending is largest, and in those countries appears to relate as much to private as public spending. In the USA, for example, Medicare and Medicaid make up about 15% of federal outlays and the desire to control them partly reflects concerns about the fiscal deficit and taxation. However, much of the rhetoric about the need to control health spending refers to the burden on businesses and households of financing private health insurance and out-of-pocket expenses, not the taxes required to fund Medicare and Medicaid. As in other countries, it is the level of overall health spending as a share of national income, rather than simply the public element, which is seen as the key indicator of the size of the affordability "problem".

The notion of the "affordability" of a health care system suggests that there is some threshold level of resources committed, some share of national income devoted to health care, beyond which societies will not wish to go. This is not to say that such a threshold level would be uniform across countries: it may depend on a variety of socio-economic and cultural factors. However, the underlying idea is that many OECD countries are reacting as if they have approached, or in some cases exceeded, their threshold.[2] At first blush, such a notion sits uneasily with the economist's approach to resource allocation, which frames the problem in terms of devoting resources to each activity until marginal (social) cost equals marginal (social) benefit. This

[2] For example, the Canadaian contribution to the OECD Health Reform project case-studies states that a consensus has developed there that , at about 9% of GDP, health care was consuming an appropriate, if not more than appropriate, share of society's resources.

of course implies that, beyond a certain point, the benefit from allocating additional resources to a particular activity will be less than the opportunity cost - in other words, the social return is less than could be attained by devoting these extra resources to an alternative use. Whereas a "threshold" suggests a wall over which no-one will want to go, the picture conveyed by the standard economic perspective on (costs minus benefits) is of going down a hill, into a valley, a then beyond a certain point going up a hill which becomes steeper and steeper.

This difference may be more apparent than real, in that what underlies current notions of an affordability threshold for health care may reflect perceptions about where on the hill we have reached. If we have already reached the stage where increasing the level of health spending produces very little gain, then in effect the hill has become steep enough to look like a wall. Such a perception about health spending appears to be widespread, as captured (at the expense of mixing metaphors here) in Enthoven's phrase about engaging in "flat of the curve" medicine: spending more and more for little or no extra benefit. When we say that we cannot afford to spend more (or perhaps even as much) on health, what is really being said is that the opportunity cost is too high: the money would be better spent elsewhere. Views about how much we can afford to spend are thus inseparable from perceptions about the benefits from additional health spending.

At a macro level, the fact that increasing health expenditures often do not appear to be producing commensurate improvements in the health status of the population fuel concerns about poor value for money. To assess properly whether such perceptions are well-founded, however, one has to be in a position to measure the impact of medical interventions and their cost-effectiveness, to which we turn shortly. Before doing so, it is worth asking what it is about health care that would get us into a situation where marginal benefit from extra spending could be so low? This takes us right back to the now-standard analysis of the ways in which health care differs from other commodities. Imperfect information, in a variety of forms, would be generally accepted as a dominant characteristic of markets for medical care. Many economists would focus on the fact that because the marginal price of medical care at point of use is often reduced towards zero by insurance of one form or another, consumers demand care until the marginal product of additional care is nearly zero. Others place more emphasis on the role of providers, given the extent to which patients have to rely on providers in making decisions about medical care. The interaction between insurance and the advance of medical technology has recently received a good deal of attention. With extensive state involvement in financing health care, some would also point to political factors which could lead to an inflation in the size of the public element in health spending.

While we return to some of these below, this paper is not the place for a lengthy recapitulation of debates about just how "different" health care is and the implications for policy. The point to be made at this stage is that the key to understanding attitudes to health spending and to "affordability" lie in perceptions of value for money, and understanding why we are on the "flat of the curve" - if we are - takes us back to fundamental issues about the nature of the market for health care. If attitudes to "affordability" are distinctive for health spending, it is because health care *is* different.

If affordability is a matter of balancing benefits against opportunity costs, who is to assess the benefits of additional expenditure on health, estimate the costs, and decide whether it is "worth" it? With most commodities in market economies, of course, this question does not arise: if consumers are willing to pay more to consume more of a commodity, standard economic analysis assumes that their welfare must be increased thereby. The reasons why simply leaving demand for health care to be expressed in the market results in an inefficient (as well as inequitable) outcome have often been rehearsed. As a result the institutional setting in most OECD countries serves to greatly reduce the role of the market. Willingness to pay, at a societal rather than individual level, nonetheless remains central to "affordability", but is very difficult to apply as a yardstick in assessing whether increased spending on health increases overall welfare.

To take an example, Newhouse (1992) argues on the basis of willingness to pay that the welfare loss associated with increases in US health spending (rather than its level relative to elsewhere) are not as large as commonly supposed. He argues that advances in medical technology accounted for much of the increase in health spending in the USA over the past half-century or so, and that consumers may have been willing to pay for most of this increase. As evidence to support that view he points out that the real rate of increase in health costs has been similar across countries with different institutional arrangements, some making explicit budgetary decisions about the size of the health care sector and trading off medical care against other public or private goods. He also notes that the rate of increase in Health Maintenance Organisation costs in the USA has been similar to the overall medical care sector (though from a lower base). If there was a public preference for less and cheaper technology, HMOs might have offered it: "if many consumers felt that new technology wasn't worth the price, it seems odd that we do not observe some firms trying to enter and offer at least some aspects of 1960s medicine at 1960s prices" (p. 16).

Such arguments are problematic for a variety of reasons, most fundamentally because of the information problems which pervade the health care sector. One can think of supply-side reasons why 1960s technology is not on offer, going back for example to the way doctors are trained, but it would

probaby be rejected by consumers in any case as "second best" care. Consumers may in that sense be willing to pay more for "up-to-date" technology, but have to take it on trust that this technology is in fact significantly better. Even if the market provided the choice, in other words, consumers would not have the information needed to decide whether spending more to obtain current technology was "worth it". Even in the US the notion that the level of health spending can be left to the market has been rapidly losing ground, as this and other imperfections in the health care market lead to widespread disatisfaction with the escalating costs of health care being expressed through the political system.

It is through the political system that societal views about the affordability of health expenditure, willingness to devote a growing proportion of national resources to health care, ultimately make themselves felt. People's views in this respect will be affected not only by their perceptions of the quality and effectiveness of the health care available to them, but also by how they perceive and value equity in the system, the care available to others in the society. What then of the possible trade-offs between affordability and quality, effectiveness and equity with which this paper is concerned? In principle, if concerns about affordability relate to poor value for money from health spending and the response is to control expenditure by improving efficiency, then health care has to become more rather than less effective and no efficiency-equity trade off need be involved either - indeed, some would argue that a comprehensive definition of what constitutes efficiency in this area would minimise the scope for such a trade-off. In practice, however, there is every danger that policies implemented to restrain expenditure - particularly some of those justified in the name of increasing efficiency - have a detrimental impact on quality of care and on equity.

To understand how this comes about, it is necessary to examine the limitations in the ways efficiency in health care is assessed and promoted. This puts the spotlight firmly on how we measure the impact of medical interventions and their cost-effectiveness, to which we turn the next section.

3. Measuring Efficacy and Effectiveness

In the face of enormous pressure to contain health-care costs without compromising quality of or access to care, more and more emphasis is being placed on improving "value for money" in health. Increasing efficiency, at both macro-level in terms of allocation of resources across for example levels of care and micro-level in terms of getting the most out of inputs allocated, has become a primary focus of policy-makers. This has served in the first place to demonstrate the extent of the gaps in knowledge about the impact of health care interventions, without which one cannot begin to assess their

cost-effectiveness. Studies from both a medical and an economic perspective have called into question the effectiveness of some health care interventions and highlighted the fact that the efficacy of many other common medical practices remains to be scientifically demonstrated.

In studies of the impact of various procedures on patients, a substantial proportion have been deemed to be medically "inappropriate" or ineffective. For example, Chassin *et al* (1987) looked at a variety of procedures, and upon clinical review considered that between one-sixth and one-third were inappropriate. A similar pattern has been found by other researchers for other procedures, and these studies have been summarised as showing that as much as a quarter of all health services currently provided may be unnecessary.[3] A side-effect of studies by economists of the factors influencing people's decisions about health care has also been to bring this issue into focus. For example, a central finding from the large-scale RAND health insurance experiment in the USA was that a fully-insured population "spends" about 40-50% more on health care than a population with a large deductible, with no measurable improvement in their health status as a result of the additional services.[4] While the first part of this conclusion has received most attention from economists, it is perhaps the second which is most striking.

How could such a high proportion of care be found to be unnecessary or ineffective? Considering this question reveals both how little is known at scientific level about the efficacy of much of current medical practice, and how often this knowledge is not reflected in what practitioners do. Large-scale control trials offer an accepted scientific procedure for testing the efficacy of a particular treatment, technology or drug. Measuring the impact of a particular medical intervention through such control trials is an expensive and time-consuming process, however, and most of current medical practice has not been validated in that way. Even for new treatments, many studies in the medical literature draw conclusions about efficacy on the basis a small number of cases, often without a control group. Formal randomized trials are themselves exercises in statistical inference, and can face difficulties in achieving adequate statistical power because cases of specific diseases, even relatively common ones, will be geographically scattered. Even when such studies have been carried out, as Phelps (1992) emphasizes, there are difficulties in diffusing among practitioners what may be learnt from the scientific literature about efficacy, and we return to the consequences of this below.

---

[3] See Borowitz and Sheldon (1993) p. 201.

[4] See Manning et al (1987).

Assessing efficacy and then cost-effectiveness requires that health outcomes be measured. In control trials, the impact of a particular treatment, technology or drug in treating a specific condition can be assessed in terms of the symptoms and limitations associated with that condition. If the condition is life-threatening organ failure, for example, the outcome measure employed may be the expected increase in life expectancy from alternative treatments, whereas for a condition such as arthritis outcomes may be assessed in terms of improvements in functioning. Given information about the costs of the new procedure compared with those of current practice, the results of such trials can be then be used to assess the *cost-effectiveness* of an "innovation" in health care. From an economist's perspective, such cost-effectiveness analysis can be seen as an invaluable aid to increasing technical efficiency - getting the most in terms of outputs from a given input of resources. If we see that a new way of treating a particular condition can produce greater improvements in health outcomes at the same cost as current methods of treatment, or the same improvements at lower cost, then substituting the new procedure will improve technical efficiency.

In practice, of course, things are not always so simple and cost-effectiveness analysis contains hidden pitfalls. To take one example, often the new procedure being assessed is expected to yield greater improvements in health, but will also cost more, than current practice. In measuring the full opportunity cost of implementing the new procedure, one should then also take into account where the additional resources are to come from and what must be given up.[5] However, the results of cost effectiveness studies are generally framed in terms of the ratio of "outputs" to costs: if the new procedure produces greater benefits per $ spent than the existing one (or alternative innovations) then it is deemed to be more efficient. This involves the (usually implicit) assumption that the additional resources will come from other unidentified areas which have rates of return no greater than the existing programme which is being used as the point of reference. Given the act of faith about the equalisation of technical efficiency across areas within the health sector, or - since the additional resources will in fact often have to be sought outside the health sector - between health and other uses, it is perhaps unsurprising that the assumption generally remains implicit.

Assessing technical efficiency in the treatment of a specific condition is clearly a limited though valuable exercise. In assessing the way resources are allocated within health care more broadly, we need to be able to compare the "value" of different interventions across the whole range of treatable conditions - how does a more effective way of treating arthritis compare with a new surgical procedure for heart disease, for example? To do so we first

---

[5] See Birch and Gafni (1992).

need a measuring rod for outcomes that can be applied across the range, before - ideally - assigning monetary values and comparing these outputs/benefits with costs. Here economists, seeking to apply standard techniques to valuation of costs and benefits, have made a major contribution. Quality Adjusted Life Years (QALYs) offer one way of combining the two key dimensions of health care output, quantity and quality: the expected increase in life expectancy and the expected improvement in quality of life throughout the patient's lifetime. The QALY concept combine these dimensions in a single quality-weighted measure of life expectancy. (Other measures in use in the literature such as Healthy Life Years and Disability Adjusted Life Years are intended to serve the same purpose).

Many issues arise in constructing and applying QALY-type measures, which have been the focus of a mini-literature all of its own and which can only be briefly touched on here. The way in which expected years of life are "quality-adjusted" is of course central, and relies for the most part on the evaluations of patients or survey respondents. The weighting or ranking of different health states is therefore intended to represent patients' (or potential patients') preferences. The extent to which QALYs as currently constructed are likely to do so in practice is contested, since there are many difficulties in trying to elicit health state valuations - there are different methods which may give different results, each method may not be reliable (i.e., may not consistently give the the same results), and it is not even clear precisely who's valuations one wants to incorporate.[6] However, if one takes it that QALYs do indeed represent preferences they can be used in Cost-Utility Analysis. This is a special form of cost-effectiveness analysis in which the outcome measure is QALYs gained (or a related measure), allowing the increase in QALYs from different innovations (across various programmes or treatable conditions) to be compared with the cost, and generally presenting results in terms of cost per QALY gained. If faced with the problem of how to allocate a given health care budget, this provides a basis for selection of programmes, with priority being given to those with the greatest "rate of return", i.e. lowest cost per QALY gained.

However, as in the case of cost-effectiveness analysis, the extension to cases where additional resources may be required is questionable, since the opportunity cost of these resources has not itself been measured.[7] Cost-utility analysis does not therefore help with the central question discussed in Section 2, that is how much resources should be devoted to health care.

---

[6] On the problems in measuring preferences see for example Loomes and McKenzie (1989), They also point out that there may be substantial differences between the valuations of random samples of the population, patients experiencing the condition in question, or health service professionals.

[7] See Weisbrod (1992), Birch and Gafni (1992).

Some cost-utility studies talk in terms of a cut-off, an acceptable level of cost per QALY gained, determining whether a new programme is worth financing. Such rules of thumb are not soundly-based, since in order to properly evaluate whether additional resources should be allocated to health care we have to be able to *value* the output in monetary terms, compare *that* with the cost, and assess whether this exceeds the return available from allocating the resources elsewhere. In other words, a fully-fledged cost-benefit analysis is required. Carrying this out runs into the extremely knotty problem of how to value health outputs, improvements in life expectancy and "quality of life".

The difficulties involved are illustrated by the extensive literature on placing a monetary value on human life.[8] Willingness to pay, on which the economist generally falls back as a standard for assigning value, is of limited use here. It is difficult to measure what people would be willing to pay for an extra year of life, though inferences can be made from the choices they make about, for example, spending on improved safety and reducing risks. More fundamentally, using willingness to pay as the measure assumes that the existing distribution of ability to pay - resources - is an acceptable basis for making decisions about health care. As Weisbrod (1991) puts it, while a person's willingness to pay is bounded by their income and wealth, there is no such bound if a social judgement is reached that the person has a property right to certain types of medical care. In that case, the basis for valuing output is not the maximum the person would pay for the treatment in question but the minimum he or she would accept to give it up - and in the case of a life-saving treatment no amount might be large enough. Societal choices therefore have to be made about precisely what forms of care are indeed considered, in effect, property rights of which no one should be deprived. Alan Williams (1977) makes the point as follows:

"If society has explicitly rejected willingness and ability to pay as an appropriate basis for rationing out health care, the problem of vindication surely shifts onto those who persist in using that criterion" [9].

Cost-utility analysis, which has a more limited objective, can be carried out without having to face these value-laden issues, and is for that reason much more widely practiced in the health care area. Many pitfalls remain in using the results of such studies to compare different types of health expenditure. For example, one sees "league tables" bringing together the results of different studies to rank various potential expenditure programmes on the basis of the ratio of incremental cost to incremental QALYs gained. However, these can be quite misleading if taken to mean that the "innovations" with the lowest cost per QALY gained should be funded first,

---

[8] See for example the valuable survey by Viscusi (1992).

[9] Manuscritp, quoted in Wagstaff (1991) p. 23.

since the incremental benefits of each programme are measured relative to *different* alternative programmes.[10] A proper economic evaluation would instead use the same point of comparison to assess the oportunity cost of different programmes, namely the highest-valued alternative use. Further, the ratio of (incremental or total) cost and "output" of different programmes may not be independent of the size of the programme itself - there may be economies of scale up to a certain point, beyond which there are constant or diminishing returns. Simply talking in terms of "a" rate of return from spending on a particular programme can thus be quite misleading. This calls into question, for example, the way the World Bank in the recent World Development Report on "Investing in Health" ranks different interventions by the ratio of cost to gains in terms of Disability Adjusted Life Years, showing which cost $1 per DALY gained, which $10 per DALY gained etc. Total cost varies greatly across interventions, but seems to be taken as pre-determined, and there is no discussion of whether it represents the point where cost per DALY gained *for that programme* is at a minimum.

Both the way QALYs (or related measures) are currently constructed and the way they are often applied are thus open to debate and criticism on a variety of fronts. Some, particularly non-economists, would question the reductionism implied in combining length and quality of life into a single measure.[11] Some economists, on the other hand, stress the limitations of cost-effectiveness and cost-utility analysis in assessing efficiency compared with full-blown cost-benefit analysis.[12] However, economic evaluation in general, and QALYs in particular, are designed to aid policy-makers in making decisions which will get made anyway, and a degree of pragmatism may be appropriate. Presenting and assessing results purely in the form of QALYs can certainly present the decision maker with a "black box", masking important information and restricting the scope for assessment of the underlying valuations. As those involved in the development of QALYs would emphasize, however, a good QALY analysis should show all the results first in disaggregated form, before combining them using values gathered from appropriate individuals, followed by a thorough analysis of the sensitivity

---

[10] See Birch and Gafni (1992). For example, moving to a new procedure for treating heart disease might be seen to yield 3 additional QALYs per patient for every extra $ spent, while a more expensive way of treating arthritis might only yield 1 additional QALY for each extra $ spent. But it could be that *total* spending on treating heart disease was very much less efficient (in terms of QALYs produced) than spending on those suffering from arthritis, so that shifting resources from the former to the latter would maximise the total output of QALYs.

[11] Cox et al (1992).

[12] See Birch and Gafni (1992).

of the results, the goal being simplicity and transparency.[13] This poses a challenge both for policy-makers and for researchers trying to provide them with information, since simplicity and transparency may conflict in practice and a balance has to be found. Nonetheless, presenting disaggregated information on quality and length of life but also attempting to combine them, rather than leaving all the hard questions to be faced, must help decision makers, and doing so on the basis of patients' or the public's valuations rather than some arbitrary weighting scheme, whatever the difficulties, seems the best route to follow.[14]

Before turning in the next section to how these measures are best applied as an aid to improving efficiency, it is worth considering the implications their use may have for the other central goal, of equity. Some concern has been expressed about the implications of basing resource allocation decisions on QALYs, not only by critics but by some who (otherwise) defend their use. The notion basic to the use of QALYs, that an extra year of healthy life is of equal value no matter to whom it accrues, can be seen as incorporating a kind of equity, likely in practice to lead to a more equitable allocation of resources across for example income groups than at present. However, some might wish for preferential treatment of the less-well off in allocating resources rather than simply maximization of total QALYs gained, on a variety of grounds. If society in effect attaches different values to an extra year of life going to different people - the young versus the old or the poor versus the rich, for example - one possibility is to incorporate this in the QALY approach by weighting differently the healthy life years gained depending on the type of person concerned. The goal then becomes the maximization of a weighted rather than an unweighted sum of QALYs, and, it is argued, equity concerns are embodied in an integrated way.[15]

How the goal of equity itself is to be interpreted in the health care context has of course been hotly debated, with "equal access for equal need", "equal treatment for equal need", and equality in health outcomes being prominent candidates. However, Wagstaff (1991) argues that the weighted QALY approach does not provide a satisfactory way to capture concerns about equity under any of these definitions, and proposes an alterrnative. This involves employing a social welfare function which explicitly incorporates an aversion to inequality in health, but permits some trade-off between inequality and efficiency. The allocation of resources would then be based on both efficiency and equity goals, taking into account how far society is prepared to accept

---

[13] See for example Torrance (1992).

[14] See for example Williams' (1992) critical comments on what he terms the "arbitrary all-purpose integer scale" suggested by Cox et al (1992).

[15] See for example Culyer (1990).

a lower per capita health status in order to achieve greater equality in health outcomes. (Welfare maximization along these lines would, he concludes, in fact imply equal treatment for equal need).

While Wagstaff makes clear that the goal of perfect equality of health outcomes is not a particularly sensible one, it is the implications of seeking to *reduce* such inequalities which he explores. It may not be the level of inequality in health outcomes which lies at the heart of concerns about equity in health and health care, though, but rather those inequalities which are a result of underlying socio-economic structures and inequalities. It is not the fact that people life lives of varying length and quality which is "unfair", from this perspective, but the fact that some people are placed by society in positions which sytematically disadvantage them in terms of health outcomes. One might then wish to give preferential treatment to the less well off in allocating health care resources in order to compensate, to some extent, for these disadvantages. Such concerns could be captured by applying different weights to different groups in applying the social welfare function approach, rather than in simply weighting QALYs gained differently. An alternative would be to present results using both unweighted and weighted QALYs, so that the impact of differential weighting could be clearly seen. Seeking to incorporate equity considerations into the outcome measure tries to do too much, further exacerbating the danger that the results from the "black box" will be misunderstood and incorrectly used. Keeping efficiency and equity considerations separate and allowing for the possibility of explicitly trading off one against another seems far preferable. While this could be seen as allowing economists to push equity considerations to one side, it should instead serve as a constant reminder to them of the fundamental result in welfare economics that "inefficient" outcomes are sometimes socially preferable.

4. Improving Efficacy and Cost-Effectiveness

It is not necessary to elaborate further on the difficulties which arise in measuring health care outputs and assessing the efficiency of different programmes and procedures. Perhaps the most daunting is simply the high cost in information terms of measuring the various dimensions of these outputs and operationalizing concepts such as QALYs. As a result, only a small proportion of the thousands of known medical procedures has been analysed systematically. What should be emphasised, though, is the progress which has been made, and how this has increased our understanding of the complex issues involved which have to be tackled rather than avoided. This illustrated by the unsatisfactory nature of some widely-used short-cuts for measuring output and productivity. For example, output of the health services is often measured in terms of hospital admissions, in-patient days, discharges, consultations etc., and costs in terms of cost per hospital day, per case or per procedure. These tell

us a great deal about the intermediate output of the health services but little or nothing about the impact of these on health status and therefore the final output which can provide the only adequate basis for measuring productivity and cost-effectiveness. A decline in hospital costs per case can be seen as an increase in productivity, for example, but could be brought about mostly by hospitals discharging patients "quicker but sicker". The dangers inherent in relying on such measures - and in rewarding providers on that basis - are clear.

If used properly, cost-effectiveness and cost-utility anaysis provide an essential tool for management and policy-makers in seeking to improve the efficiency of the health services. It does not provide a cookbook or recipe to apply, but this is a reflection of the complexity of the problem rather than the failure of the approach. Indeed, given the pressures on health budgets, some analysis of costs relative to output will inevitably underlie critical decisions about how health budgets are spent - the only question is whether it will be done well or badly. Formal economic evaluation using current methodologies may only properly applicable within fairly restrictive boundaries, but can make a major contribution to promoting more systematic decision-making.[16] We have seen that it is least problematic from a methodological point of view in deciding how best to allocate a fixed budget over a narrow area or how to minimize the costs of dealing with a specific illness, for example deciding what drug to use in treating a particular condition. The contribution which such a focus can make to controlling expenditure has been amply demonstrated by, for example, the savings generated by the increasing use of generic drugs.

Over broader areas, such as how much to spend in treating different diseases, many more difficulties and uncertainties arise. Researchers most closely involved in the development and application of economic evaluation tend to emphasise these difficulties, and can be rather modest in the claims made on its behalf.[17] The desire to incorporate cost-effectiveness more systematically than heretofore into decision-making is nonetheless illustrated by the increasing pressure for the use of formal economic evaluation, often including QALY-based rankings, at official level. Since governments currently insist that drugs are more evaluated for safety and efficacy than devices or medical procedures, it is there that economic evaluation is likely to be most easily built in to official decision-making (as illustrated by recent Canadian and Australian experience). The use of such methods in evaluating other aspects of medical technology requiring capital investment which is subject

---

[16] The rapidly increasing volume of economic evaluation studies being produced can be seen from the bibliography recently compiled by Backhouse et al (1992).

[17] See for example Drummond (1992).

to government approval is likely to be next. While cautioning against over-hasty adoption and misuse of methods of economic evaluation, economists have to be aware of the needs of policy-makers and the limited information on which decisions are in fact taken, and accept the responsibility for enhancing that information as best they can. Since the confidence intervals around such results will be wide, perhaps the most useful goal at this stage would be to try to distinguish "very good buys" from "very bad buys" in a robust fashion. Governments and societies will still have to struggle with the question of how much to spend on health care versus other goods and services, but given the extent to which equity considerations affect that decision the limited scope for application of formal cost-benefit analysis there may not be as great a limitation as it seems.

Apart from playing an important part in the assessment of new drugs and devices, economic evaluation could have even more potential in reducing the extent of inappropriate utilization, which as described earlier appears to be considerable. One route to reducing inappropriate utilization and improving the overall effectiveness of health care is through reducing the extent of variations in treatment. The extent of the variation in medical practice across OECD countries and across regions within these countries is by now well-documented, one consequence of limited knowledge about efficacy, at scientific and even more so at practitioner level. The analysis of such variations goes back to the 1930s, when it was shown that the rate at which British school children had their tonsils removed varied greatly from one region to another. Considerable variation has subsequently been found for various surgical and other procedures, both within and across countries. In addition to tonsillectomy, surgery for low back injuries and coronary artery bypass are examples of procedures where high levels of variation have been documented in various countries and regions. The extent of variation in the frequency of Caesarian section operations in childbirth has also been frequently noted, as have major differences in drug prescribing patterns.[18]

This variation does not appear to be simply a reflection of differences in age profiles or illness patterns of the populations in question. Nor are they in general explicable in terms of differences in economic incentives, insurance coverage, provider payment systems etc. across areas, since variability has been found in countries with very different health care systems such as the USA, Canada, Britain and Sweden. Finally, it is not in general the case that the observed variation simply reflects different choices among alternative treatments for the same disease. In fact, there are often major differences in the overall rates of treatment of diseases, with the use of interventions that could serve as medical substitutes positively rather

[18] See Phelps (1992) Table 1 p. 25. See also McPherson (1989) on cross-country differences in practice.

than negatively correlated. Reducing the extent of variations in treatment could result in major savings. For example, Phelps and Parente (1990) estimate that for the USA, the aggregate welfare loss from variation in hospitalization rates across regions exceeded $7 billion per annum. They remark that even if some of the observed variation is in fact desirable, the welfare loss produced by the remainder appears to be of the same general magnitude as that commonly attributed to moral hazard arising from the impact of insurance on incentives to use medical care. Reducing variations in practice is very different to observing them, however. Indeed the persistence of these variations is if anything more striking than their existence: Phelps notes that the extent of variation in the use of tonsillectomy in studies carried out in the late 1980s closely matched that found in the 1930s. Unlike other sectors, in health care a superior "technology" or treatment strategy does not appear to drive out other treatments over time.

This partly reflects the limited knowledge about efficacy in many areas which has already been mentioned, but also arises because of the way in which scientific knowledge is diffused to practitioners. Most doctors provide treatment for a very large number of different diseases, and no individual could hope to keep up in different areas with the latest studies reported in scientific journals. Most therefore rely largely on what their peers in a particular locality are doing and on their own experience. Local "testing" of a new therapy by one or two doctors on the basis of a small number of cases may therefore have a significant impact on whether it is adopted in that community. Thus even when there is a consensus at scientific level about the efficacy of a treatment or when it is appropriate, this does not automatically feed through to local practice: although the view has prevailed for many years at specialist/scientific level that widespread use of tonsillectomy is undesirable, the substantial variations in its use have persisted.[19]

One way to try to reduce variations in practice which is currently receiving a good deal of attention, notably in the USA and Canada, is the development of "protocols" or practice guidelines which set out the appropriate treatment strategy in different circumstances, based on the best scientific evidence about efficacy.[20] The difficulty of designing and agreeing on such protocols is not to be down-played, since the problems facing a practitioner - the range of situations, the mix of symptoms and the variety of treatment options - are often too complex to be readily amenable to standardization. In addition, very often not enough is known at scientific level about efficacy. What is worth stressing here, though, is that what is

---

[19] See Phelps (1992) p. 35.

[20] The US federal government is currently investing about $100 million per year in research on this topic, as documented in the US case-study for the OECD Health Reform project.

known about cost-effectiveness could also be incorporated into the guidelines. On this basis, claim its proponents, practice guidelines offer a unique method of cost containment in that they have the potential to control costs and improve the quality at the same time. [21] This obviously adds another layer of complexity to the production of guidelines, and requires close co-operation across the disciplines. Economists will also be particularly interested in exploring ways of providing doctors with an incentive to follow the guidelines, since the available evidence suggests that the production of guidelines in itself will not be sufficient to alter clinical practice.[22] Increasing both the rate of production and dissemination of information about the efficacy of medical treatments offers a relatively inexpensive way of improving the effectiveness of the health care system. Formulating practice guidelines which also take cost-effectiveness into account appears a logical extension, having the potential to improve effectiveness while controlling costs.

5. Insurance, Quality and Cost-Effectiveness

In OECD countries, the risks of incurring substantial expenditure on health care are commonly spread by insurance, with institutional structures varying greatly not only in the public/private mix but also within each of these elements. Such pooling of risks partly reflects concerns about equity and access to health care, but leaving aside such considerations efficiency arguments still call for insurance. As in other areas, insurance provides a response to uncertainty about the future, in this case about whether ill-health will strike. Debates about the extent to which the private market can supply health insurance efficiently revolve around the problems of adverse selection and moral hazard. Once again it is not intended to rehearse these arguments in depth here: we concentrate on the way insurance affects the relationships between affordability and quality, effectiveness and equity which are the concern of this paper.

The effects on consumer behaviour of the moral hazard associated with insurance have been a recurring theme of health economists, the fact that the marginal price at point of use is much lower than marginal cost being seen as leading to overconsumption of health care. Particularly where health insurance is itself subsidised, this has frequently been identified as a major factor in fuelling the increase in overall health expenditure (see Pauly 1986). One reaction of both private insurance and many public systems to spiralling costs has been to increase the extent of coinsurance - through the increased use of

---

[21] Mark Chassin, as quoted in Borowitz and Sheldon (1993) p. 202.

[22] See Borowitz and Sheldon (1993).

deductibles and co-payments by insurers, and through public systems introducing or raising prices at point of use. Apart from equity concerns which are dealt with below, this could only be a partial solution since even with a significant coinsurance rate of, say, 20% (now typical in the USA), the patient's marginal price is still well below marginal cost. If the coinsurance rate for substantial medical bills was very much higher than that, the whole basis of risk-pooling would be undermined (leading in all probability to a demand for supplementary insurance).

In any case, on the face of it moral hazard is more plausible as an explanation for "excess" health spending at a point in time than for sustained increases over time. As Newhouse (1992) points out, although the spread of insurance coverage over the 1950-1980 period in the USA was substantial, cross-section estimates would predict that the response of demand would account for only a small proportion of the actual increase in expenditure. As we have seen, it is technological change which he sees as the main driving force behind increased expenditure. However, the minor role this assigns to insurance may not take sufficient account of the ways in which insurance can serve to encourage research and development and the introduction of new technologies. As Weisbrod (1991) argues, the interactions between insurance and technological change may be crucial, with new technologies fuelling the demand for insurance and insurance providing incentives for the development of cost-increasing technologies. When a new technology - such as the CT scanner or organ transplantation - is developed, its speed of diffusion and subsequent effect on health-care expenditure depends heavily on the extent to which its cost can be borne through the insurance system. Equally importantly, the expectation that these costs will be borne by insurance may be critical in creating an environment conducive to the long-term investment in research and development required to develop new cost-increasing technologies.[23]

Insurance can make very costly technological advances possible, but it can also enhance the profitability of expensive innovations that offer only marginal therapeutic benefits. Not only can this lead to operating on the "flat of the curve" with current technology at a point in time, but insurance contributes to the situation where more and more can be spent over time for smaller and smaller improvements in health care. "Insurance" is of course being used here to denote private or public coverage of the costs of health care, and it is not simply the presence of insurance but its form and payment structure which is important. It has been recognised since the 1970s that where public or private insurance covers the costs of health care on a retrospective basis, there is little incentive for providers to weigh costs against patient benefits. What has recently received more emphasis, though, is the effect this could have on the development and introduction of new

[23] See also Gooderis and Weisbrod (1985).

technology. Those involved in research and development will focus on developing technologies that enhance the quality of care, regardless of cost, if they are confident that providers will adopt such technology and that insurance will cover the cost. The introduction of new, costly technology may then in itself increase the demand for insurance since the consequences of being without insurance - in terms of cost of obtaining care if one becomes ill - become more serious. A welfare loss is not necessary involved, but retrospective cost reimbursement certainly contributes to a situation where new technologies can be introduced which are welfare decreasing, in the sense that the insured population is not willing to pay the real cost of developing and applying the technology.

Both public sector and private insurance have commonly reacted to spiralling health spending by moving towards prospective payment systems. Does this fundamentally change the incentives for the development and adoption of new technologies, with implications for the "quality" of care? The signal being sent to the R&D sector is no longer "Develop new technologies that enhance the quality of care, regardless of the effect on cost", but instead "Develop new technologies that reduce costs, provided that quality does not suffer 'too much'". Taking technology as given, prospective payment also provides encouragement to providers choosing between alternative treatments to use those that reduce costs, rather than improve quality regardless of cost as is the case with retrospective cost-based payment. Following the shift to prospective pricing, for example, a move back towards re-usable rather than disposable medical devices has been seen. Such changes in practice are not necessarily inefficient or detrimental to quality: the impact on costs is unambiguous while that on quality is uncertain. This is an illustration of the general point that the switch to prospective payment would be expected to change practices in a manner that has a favourable impact on costs, but "subtle or uncertain, but presumptively nonpositive, effects on the quality of care".[24]

To what extent can protecting the quality of care be left to consumers - relying on competition between providers/insurers and the market standard of *caveat emptor*? The scope for consumers to exercise "sovereignty" and respond to a perceived deterioration in quality by taking their business elsewhere obviously depends on the institutional structure, and in many countries is limited. In a more market-oriented environment where consumers are choosing between competing providers or insurers (or jointly HMOs), can the problem of their imperfect information about prices, waiting lists etc but especially about quality be overcome? Proponents of (managed) markets in health care assign a great deal of importance to improving the information available to consumers. They envisage for example the collection and publication of much

[24] Weisbrod (1991) p. 543.

more information about activity and outcomes for individual hospitals or health plans. This would be undertaken by insurers and/or regulatory authorities.

However, measuring performance and constructing "league tables" for hospitals or health plans can be highly misleading if insufficient account is taken of differences in age profile, case-mix and severity of illness of the patients treated. This is analogous to, but more serious than, the difficulties which arise in comparing school performance. There, comparisons on the basis of examination results alone can be highly misleading about effectiveness, since there can be enormous differences in the student "population". In the case of health care - for example hospitals - we need in the same way to standardise comprehensively for differences in the age, severity, case-mix etc. of the patients treated, and taking all relevant differences into account is likely to be just as difficult as in education. In addition, though, we also face major difficulties in measuring and comparing "results". Measuring differences in quality rather than comparing "throughput" is of central importance, and yet little progress has been made in developing ways of measuring the quality of output of a hospital - indeed consistent measures of hospital output itself are generally only available as a by-product of the move to prospective payment systems.

Quality in the case of health care is inherently difficult not only for individual consumers but for insurers, regulatory agencies or the state to monitor. This problem is accentuated by the fact that the incentive facing providers under prospective payment is to reduce precisely those aspects of quality which are most difficult to observe. (Thus regulators will find it much easier to monitor the physical environment of a hospital rather than actual clinical practices). The likelihood that quality of care will suffer may be greater where prospective payment is combined with competition on the supply side, whether within the public or private sectors or between these sectors, since the pressure to control costs is accentuated. It is also important to note that the response to changing financial incentives may depend on the type of institution involved - private, government, and non-profit organisations may not respond in the same way. When all three institutional forms are present in the USA there is some evidence of differences in behaviour, with non-profit providers perhaps providing a wider range of services more willing to care for the poor, but this is a contested area where much research remains to be done.

Since the shift to prospective payment changes incentives so radically, why has that alone not been enough to control costs, most obviously in the USA where it was widely adopted in the 1980s by both public and private sectors? One answer is that technology is central, that the shift to prospective payment will have the most powerful effect through its influence on R&D, and

that by its nature this takes a long time to be felt in the technologies coming "on stream" and on costs of care. For example, costly drugs recently becoming available would have been in development before prospective payment was widespread, as would costly diagnostic devices such as CT scanners and MRI devices, but the climate for investing in R&D on such technologies is now very different.However, focusing on the structure of the payment system and its effects on incentives may again not do justice to the dynamics of the interaction between technology and insurance. Simply noting that prospective payment systems have been widely adopted does not in fact tell us anything like the complete story about incentives to develop and adopt new medical technology. One also needs to know how the prospective payment rates will be adjusted over time to take into account the impact on costs - positive or negative - of new technologies, and the basis on which decisions to accept or reject new technologies will be made. This cuts both ways. On the one hand, it could be naive for investors in R&D to assume that procedures for which new cost-reducing technologies are developed will not see a cut in their reimbursement rate after a time, so the incentives to concentrate on such technologies may be weaker than they appear in a static setting. On the other, cost increasing technologies may of course still be adopted, and reflected in increased reimbursement rates, if they are seen to be "worth it" in terms of improved outcomes.

This brings us back to how the cost effectiveness of new technology is to be assessed, who is to do that assessment, and how it is to feed through to decisions about which new technologies are to be adopted and how they are to be reflected in reimbursement rates. Those developing technologies must be reasonably confident that insurance cover will include their "product" at a rate which provides a return on their investment. Prospective payment and other changes in the way insurance is provided are likely to increase uncertainty for those developing medical technologies, but the long-term impact on technology will depend on how the system is seen to work in practice and particularly how reimbursement rates are adjusted over time and how decisions about adoption of new technology are made. Here the ethical and political dimensions of health care are also crucial. Where there is a widespread view in the society that "high-quality" health care ought to be available to everyone, there will be public pressure for the state to ensure that access to costly new technologies is not confined to the wealthy.

6. Insurance, efficiency and Equity

The objective of (public or private) health insurance is to pool risks, but how should the costs of health care then be spread over the insured population? This is of course a central issue from an equity perspective. From

that perspective, the notion that the costs of providing health care should be met simply on the basis of ability to pay is often advanced. A weaker equity aim would be met by having everyone pay the same irrespective of age, gender etc. - community rating, in insurance parlance - so that there is redistribution from healthy to ill (and therefore from young to old, i.e. over the life-cycle, and perhaps also to some extent from rich to poor). On the other hand, those who advocate relying primarily on financial incentives to consumers choosing among insurers to guide the allocation of resources tend to see perfect risk-rating as ideal in promoting an efficient outcome. The way costs are spread thus has a major bearing on the way competitive insurance markets will operate, and just how "managed" managed competition might have to be in order to meet equity and efficiency goals. Focusing therefore on the situation where there are competitive insurers, to what extent could and should they rate their insured population on the basis of individual risk?

The key issue is once again the nature and consequences of information asymmetries, here between those providing the insurance and those obtaining it. Most attention has been paid to the consequences in terms of adverse selection and "cream skimming". Adverse selection may arise where the purchaser of health insurance has more information about his or her riskiness than the insurer. If the latter cannot distinguish between good and bad risks and charges the same premium to each, those who are actually low risk may find the cover unattractive so those choosing to buy insurance will be riskier than average, riskier than the insurer expected. Premiums will have to be raised to cover their health care costs, possibly resulting in a spiral of a decreasing and ever-riskier insured population and rising premia.

For adverse selection to be a serious problem, clearly individuals must be in a position to predict their riskiness better than insurers, and must act on the basis of that information in deciding whether to buy insurance or how comprehensive the cover bought will be. The Rand Insurance Experiment showed that people were in fact good predictors of their own future medical expenses, and those predicting relatively high expenditure were more likely than others to take out comprehensive cover (Marquis 1992). In a simulation exercise this suggested that when insurers offered a single premium - community rating - adverse selection would be serious enough to drive out more comprehensive plans. Can insurers then risk-rate their pool of potential customers so as to overcome this problem? The simulation exercise also showed that when insurers varied premia according to the age and sex of the family head, adverse selection was reduced but not eliminated.

One response is that insurers could and would obtain and use more information about the insured to improve their rating of riskiness, and thus further reduce the scope for adverse selection. Sloan (1992) for example points out that insurers could require policyholders to answer questions about

health status, or even insist on a medical examination. Why then do insurers put so little effort into improving their risk-rating, often setting premia without reference to even the basic demographic data which they already have available? Obtaining extra information - particularly through regular examinations - could be costly, but that does not explain why existing information is often not used.

There is a more fundamental issue, though. If insurers could and did finely risk-rate the premia charged to individuals, would that be a satisfactory situation? With "perfect" risk-rating, individuals would pay the actuarial value of their losses in premiums. High-risk individuals, for example the elderly and the poor, would face higher premia than the young and the better-off. Although their need for insurance is greater, those at high risk are less likely to purchase cover or, if some cover is mandatory, likely to purchase less comprehensive cover. Perfect risk rating, while actuarially fair and solving the problem of adverse selection, would thus be regarded by most as unacceptable from an equity perspective precisely because of the extent to which it limits the pooling of risks.

With pooling of risks across risk groups, however, insurers have an incentive for "cream skimming" or preferred risk selection - they may be able to distinguish and confine cover to relatively low risk individuals. This is in a sense the obverse of adverse selection, in that it is now the insurer rather than the individual taking advantage of information about riskiness. Even if insurers are obliged by regulation to have open enrollment, they can in various ways seek to attract low-risk individuals and discourage high-risk ones. Rather than competing primarily by being more efficient, insurers can concentrate much of their effort on selection of their insured population. Risk selection by insurers can thus undermine attempts to retain the benefits of competition between insurers and avoid adverse selection but still spread risks over the population.

The extent of regulation and state involvement in managing competition which might be necessary in these circumstances if equity and efficiency goals are to be attained is illustrated by both the Dutch reforms flowing from the Dekker Commission proposals and the variants of managed competition currently under consideration in the USA. The Dutch reforms, for example, are to involve *inter alia* a central fund into which all individuals pay income-related premia, out of which risk-related subsidies covering most of the cost of health insurance will be paid directly to the insurer of their choice. The remainder will be covered by the premium paid directly by individuals to insurers, which has to be identical for all those covered by a particular insurer but can vary across insurers. Insurers will have to have regular periods of open enrollment and antitrust policy will attempt to combat anticompetitive behaviour by insurers or providers. Among the many conditions

necessary for success, it is worth highlighting here the crucial role played by the rating of risk in setting the differential subsidies to be paid from the central fund to insurers, which will account for most of their income. Suppose the rates paid do not accurately reflect the costs which insurers will incur, and insurers are able to distinguish groups where the rates are likely to exceed costs from those where they are likely to fall short of them. The primary incentive for insurers is once again not to reduce costs by greater efficiency and compete for business by quoting a lower premium than competitors, but to "cream skim" - now meaning trying to select not necessarily low-risk individuals, but those from groups which even if high-risk have a high subsidy relative to expected costs. The best way to avoid this is obviously to have rates as closely aligned as possible with actual risks,[25] but this obviously assumes that risks can be predicted as well by the health management authority as by insurers. (Reinhardt (1993) for example draws attention in the context of the US debates to the difficulties experienced in classifying the elderly by risk in administering Medicare).

More generally, the Dutch case is a good example of the trade-offs which inevitably arise if the consumer is to act as the fulcrum of the system in promoting efficiency. Can the consumer be provided with sufficient incentive to "shop around" and thus promote efficiency among insurers and providers, while retaining "enough" pooling of the risks of ill-health across the population? Of course, what is regarded as "enough" pooling of the costs of health care may vary from one society to another. In the Dutch case, however, the degree of pooling seen as desirable is such as to leave only a small proportion of the costs of care payable directly by the individual, which may leave an insufficient incentive for people to very actively seek the best value in insurance.

The trade-off between having an acceptable degree of pooling and providing incentives to patients is central to the major issue facing health policy-makers, which is (as Reinhardt (1989) puts it) whether adequate control over resource allocation to and within health care is best achieved by letting consumers (patients) discipline providers competing against each other on the basis of quality and price, or through regulatory controls on the supply side. A central element in the former approach is "the emancipation of patients into cost-conscious, prudent purchasers, by forcing them to bear substantially more cost sharing at point of service".[26] Although the merits of this approach are most often debated in the US context, it is striking how many OECD countries have included among recent health care reforms (implemented or proposed) an

---

[25] See van de Ven (1991) p. 438.

[26] Reinhardt (1989) p. 339.

increase in out-of-pocket charges to patients.[27]

From an equity perspective, this is a cause for serious concern both in terms of its impact on financing and on utilisation. A recent collaborative cross-country study on equity in the financing and delivery of health care in 12 European and North American countries, sponsored by the Commission of the European Union, has measured the progressivity or regressivity of different sources of financing.[28] Applying a common methodology, the results showed that, across the various countries with widely differing institutional structures, out-of-pocket payments were clearly the most regressive form of financing. This was the case even in those countries - like Ireland - where those on low incomes receive concessionary treatment in the application of such charges or payments. In terms of equity in access to and use of services, charges will have a greater impact on the less well-off, and it may well be "necessary" rather than "unnecessary" utilisation which is discouraged. Given the limited information available to patients relative to providers, and the stressful circumstances in which they often have to make health care choices, a major question-marks rests over the assumption that health care consumers can provide the necessary demand-side impetus to efficiency if only it is made "worth their while". On this basis, it seems reasonable that the burden of proof rest with those arguing in favour of greater reliance on generalised out-of-pocket payments as a way to promote efficiency. (Charges specifically designed to discourage inappropriate use of higher-level facilities, for example hospital out-patient departments instead of a general practitioner, are more easily justified). [29]

An intermediate route which need not involve substantial out-of-pocket expenditure but still depends primarily on demand-side incentives is to rely on the consumer's choice between insurers (or HMOs) to provide the motive power to the system. This is of course the thrust of the Dutch reforms, and of some of the many variants of "managed competition" currently being debated in the USA. With sufficient subsidisation of high-risk patients this may produce an acceptable level of risk-sharing, though the danger that this will lead insurers to concentrate on cream-skimming is not to be under-estimated. Returning to an earlier theme, the problem of protecting quality also looms large: if insurers/HMOs are under pressure to reduce costs in order to make a profit while offering competitive premia levels, quality of care - particularly those aspects which are not readily observable by patients - may

---

[27] Countries where public health service charges have recently been introduced or substantially increased include Finland, New Zealand, France, Ireland, and Germany. Canada, by contrast, has eliminated charges entirely.

[28] See van Doorslaer, Wagstaff and Rutten (1992).

[29] For an assessment of the efficiency and equity arguments for and against charging for public health care, see Nolan (1993).

suffer.

In a situation where *caveat emptor* cannot be the guiding principle and insurers may face incentives to cut corners, one is necessarily reliant to a great extent on the professional standards of health care professionals in safeguarding quality of care. However, the position in which these professionals often now find themselves must be highlighted. The notion that the patient and doctor are in a "principal-agent" relationship is a familiar one. Because of his or her professional knowledge, the doctor is in a better position than the patient to know about the likely outcomes of alternative treatment options. A doctor acting as perfect agent would serve the interests of the patient by assisting him or her to reach the decision which fully utilises that information but also reflects the patients own wishes and tastes - in effect, the decision the patient would reach if he knew as much as the doctor. Increasingly, though, doctors are being called on to act not only as protectors of the patient's interests but also protectors of the fiscal interests of public and private insurers: they are thus "double agents".[30] Indeed economists are often to the forefront in seeking to persuade doctors of the need to take on this role and its consistency with medical ethics - "it is ethical to be efficient".[31] This is altogether apart from the likelihood that doctors will take their own financial interests into account in rendering advice - and are therefore "triple agents" - which continues to sharply divide economists. This means that doctors (and other health care providers) are increasingly faced with conflicts between different roles, and expecting them to carry the full burden of protecting the quality of care in these circumstances is unreasonable.

The best way to safeguard the quality of health care is to provide more, and more reliable, information on outcomes and input-outcomes relationships to policy-makers, professionals and the public. Indeed, greater accountability to the public, greater choice and more emphasis on patients' rights as a means of quality assurance is a common theme in recent health system reviews, for example in Canada, The Netherlands, Norway, New Zealand, Sweden and the UK. Here the state must play the central role in ensuring that such information is produced. This brings us back to the issues raised in our earlier discussion about measuring outcomes and effectiveness: much remains to be done in developing the analytic techniques, but perhaps the most pressing need is for existing techniques to be applied much more widely to existing medical practise as well as potential "innovations".

---

[30] See Blomquist (1991), Weisbrod (1992).

[31] See Culyer (1992).

7. Conclusions

Widespread concerns about the "affordability" of health care reflect, we have argued, not so much worries about the proportion of national income being absorbed by health care but perceptions that this is not all money well spent. Remarkably little is known about the efficacy of much of current medical practice, and even less about its cost-effectiveness. Assessing efficacy and cost-effectiveness is so difficult because measuring the final output of the health care sector, the impact on health outcomes, is problematic. Economists have much more to contribute to the development of the analytic techniques required, but perhaps the most pressing need is for existing techniques to be applied much more widely to existing medical practise as well as potential "innovations".

When output of the health care system is properly measured, there should be no trade-off between increasing cost-effectiveness and quality of care - on the contrary, they go hand in hand and contribute to making health care "affordable". With some of the measures currently used to measure health care outputs, on the other hand, and reward systems structured on that basis, there is every danger that there will in fact be a trade-off between (apparently) improving cost-effectiveness and quality of care. On the basis of increasing knowledge about efficacy and enhanced ability to measure outputs, policies to actually improve cost-effectiveness, rather than reduce costs at the expense of quality, can be designed. In the same way, at conceptual level there should be no conflict between efficiency and equity, if efficiency is seen in terms of making the maximum contribution to raising the general health of the community. In practice, there is every danger that measures implemented to improve cost-effectiveness, narrowly defined and poorly measured, will hit hardest those most in need.

It may be in these areas - in the measurement of the outputs and assessment of cost-effectiveness, with all its problems - rather than in the grand debates about how the health services should be financed and structured, that economists and policy-makers from different countries have most to offer each other. Harking back to Keynes' notion of economists being able to function at some point in the future as technicians - like dentists, he said - it may be that as far as readily transferable knowledge is concerned we should be concentrating on how to measure and improve technical efficiency in health care. On the broader questions of overall resource allocation and equity, the choices made and the institutions in place in different countries are highly culture-specific - indeed, they tell us a great deal about the nature of the societies involved. Slow to give advice on these issues, we should also be careful about allowing our domestic agendas to be set by other's debates. Even within one's own country, a degree of modesty about the contribution which economics can make, and a greater awareness of the extent to which normative

analysis intrudes, is required. It is all too easy to mis-apply standard economic analysis in the pursuit of one's own goals, ignoring for example the fact that "inefficient" outcomes are sometimes socially preferable.

REFERENCES

Backhouse, M., R.J. Backhouse and S. Edey (1992): "Economic Evaluation Bibliography", *Health Economics*, vol. 1 (Supplement), pp. 1-236.

Birch, S. and A. Gafni (1992): "Cost-Effectiveness/Utility Analyses: Do Current Decision Rules Lead Us to Where We want to Be?" Journal of Health Economics, vol. 11, pp. 279-96.

Blomqvist, A. (1991): "The Doctor as Double Agent: Information Asymmetry, Health Insurance, and Medical Care", Journal of Health Economics, vol. 10, pp. 411-32.

Borowitz, M. and T. Sheldon (1993): "Controlling Health Care: From Economic Incentives to Micro-Clinical Regulation", *Health Economics*, vol. 2, pp. 201-204.

Chassin, M. et al, (1987): "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services?", *Journal of the American Medical Association*, November 13, 258, pp. 293-313.

Cox, D.R., et al (1992): "Quality of Life Assessment: Can We Keep it Simple?", Journal of the Royal Statistical Society, Series A, vol. 155, Part 3, pp. 353-93.

Culyer, A.J. (1992): "The Morality of Efficiancy in Health Care - Some Uncomfortable Implications", *Health Economics*, vol. 1, pp. 7-18.

Drummond, M. (1992): "Cost-Effectiveness Guidelines for Reimbursement of Pharmaceuticals: Is Economic Evaluation Ready for its Enhanced Status?", *Health Economics*, vol. 1, pp. 85-92.

Gooderis, J. and B. Weisbrod (1985): "What We Don't Know About Why Health Expenditures Have Soared: Interaction of Insurance and Technology", Mount Sinai Journal of Medicine, Vol. 52, pp. 685-91.

Hurst, J. (1992): The Reform of Health Care: A Comparative Analysis of Seven OECD Countries, OECD: Paris.

Loomes, G. and L. McKenzie (1989): "The Use of QALYs in Health Care Decision Making", *Social Science and Medicine*, vol. 28, pp. 299-308.

Manning, W.G. et al, (1987): "Health Insurance and the Demand for Medical Care: Results from a Randomized Experiment", American Economic Review, vol. 77, pp. 251-77.

Marquis, M.S. (1992): "Adverse Selection with a Multiple Choice among Health Insurance Plans: A Simulation Analysis", Journal of Health Economics, vol. 11, pp. 129-51.

Newhouse, J.P. (1992): "Medical Care Costs: How Much Welfare Loss?", Journal of Economic Perspectives, Vol. 6, pp. 3-21.

Nolan, B. (1993): *Charging for Health Care in Ireland: How and Why?*, Policy Research Series Paper, Economic and Social Research Institute: Dublin.

OECD (1993): Health Systems Statistics (Title?), 2 vols., OECD: Paris.

Pauly, M. (1986): "Taxation, Health Insurance and Market Failure in the Medical Economy", *Journal of Economic Literature*, vol. 24, pp. 629-75.

Phelps, C.E. (1992): "Diffusion of Information in Medical Care", Journal of Economic Perspectives, Vol. 6, pp. 23-42.

Phelps, C.E. and S. Parente (1990): "Priority Setting for Medical Technology and Medical Practice Assessment", *Medical Care*, August, vol. 28, pp. 703-23.

Reinhardt, U. (1989): "Economists and Health Care: Saviours or Bulls in a China Shop?", American Economic Review, vol.

Reinhardt, U. (1993): "Comment on the Jackson Hole Initiatives for a Twenty-First Century American Health Care System", *Health Economics*, vol. 2, pp. 7-14.

Sloan, F. (1992): "Adverse Selection: Does It Preclude a Competitive Health Insurance Market?", Journal of Health Economics, vol. 11, pp. 353-56.

Torrance, G. (1992): "Discussion on Paper by Cox et al", Journal of the Royal Statistical Society, Series A, vol. 155, Part 3, pp. 388-9.

van Doorslaer, E., A. Wagstaff and F. Rutten eds. (1993): *Equity in the Finance and Dedlivery of Health Care: An International Perspective*, Oxford: Oxford University Press.

Viscusi, W. K. (1992): "The Value of Risks to Life and Health", Journal of Economic Literature, vol. XXXI, pp. 1912-46.

Wagstaff, A. (1991): "QALYs and the Equity-Efficiency Trade-Off", *Journal of Health Economics*, vol. 10, pp. 21-41.

Weisbrod, B.A. (1991): "The Health Care Quadrilemma: An Essay on Technological

Change, Insurance, Quality of Care and Cost Containment", Journal of Economic Literature, Vol. XXIX, pp. 523-52.

Weisbrod, B.A. (1992): "Productivity and Incentives in the Medical Care Sector", Scandinavian Journal of Economics, Vol. 94, Supplement, pp. 131-145.

Williams, A. (1992): "Discussion on Paper by Cox et al", Journal of the Royal Statistical Society, Series A, vol. 155, Part 3, pp. 390.

World Bank (1994): *Investing in Health: World Development Report 1994*, The World Bank: Washington D.C.