

MEASUREMENT OF HEALTH



MEASUREMENT OF HEALTH



Office of Health Economics
12 Whitehall London SW1A 2DY

No 77 in a series of papers on current health problems published by the Office of Health Economics. Copies are available at £1.00
For previous papers see pages 26 and 27.

© June 1985. Office of Health Economics.

ISSN 0473 8837

Cover illustration, by courtesy of the Mary Evans Picture Library

This paper was written by George Teeling Smith

Office of Health Economics

The Office of Health Economics was founded in 1962 by the Association of the British Pharmaceutical Industry. Its terms of reference are:

To undertake research on the economic aspects of medical care.

To investigate other health and social problems.

To collect data from other countries.

To publish results, data and conclusions relevant to the above.

The Office of Health Economics welcomes financial support and discussions on research problems with any persons or bodies interested in its work.

Introduction

Expenditure on health care is continuing to rise in all Western countries, both in total and as a percentage of gross national product (Table 1). This has underlined the political importance of demonstrating that this expenditure is giving value for money, both in specific instances and in its totality.

In other types of human activity, such as the production of goods or the provision of private services, the success of a venture can be measured by its profitability. The more the public want the goods and services, the more they will pay for them, and the greater will be the rewards for the producers and the providers. But in all aspects of welfare – health, education and the social services – profit has very often been eliminated as a measure of effectiveness and efficiency. Hence there is a need, in the welfare services, to fall back on other measures of success. And, although it is relatively easy to measure the costs of the services provided, it is much more difficult to measure their outcome in quantitative terms.

This report deals with recent developments in relation to health care which have been concerned to provide more sophisticated measurements of the outcome of treatment. This means, in effect, developing methods of measuring 'health' itself, in terms of the length and quality of life for the individuals in the community.

Table 1 Health spending as a percentage of GDP.

	1960	1980	1982
Australia	5.1	7.0	7.6
Canada	5.5	7.3	8.0
Finland	4.2	6.5	6.8
France	4.3	8.5	9.3
German Federal Republic	4.8	8.1	8.2
Greece	2.9	4.3	4.4
Italy	3.9	6.8	7.0
Netherlands	4.1	8.6	9.1
Norway	3.3	6.8	6.8
Sweden	4.7	9.5	9.8
Switzerland	4.1	7.2	7.8
United Kingdom	3.9	5.7	5.9
United States of America	5.3	9.6	10.6
OECD Average	4.3	7.4	7.8

Source OECD.

Historical background

Broadly speaking, the development of most effective medical treatments only dates back fifty years or so. During the 19th century, and the early decades of the 20th, very few treatments of unquestionable value were available. Doctors and nurses provided palliative care in the hope that nature would effect a cure. There were exceptions, such as the use of digitalis for heart disease, or the repair of a simple fracture. But the vast majority of the common serious illnesses, such as tuberculosis, pneumonia or diabetes, were without effective treatments in the early 1900s.

Almost as soon as more successful therapies became available in the 1940s and 1950s, doctors realised the importance of evaluating their effectiveness. This led to the general introduction of the so-called 'clinical trial'. Systematic measures were introduced to assess the efficacy, initially of new medicines and later of surgical operations. This is now very often done in controlled randomised clinical trials, which compare the new treatment under evaluation either with a placebo or the previous standard treatment. If possible, this is done on a 'double-blind' basis, when neither the patient nor the doctor knows at the time whether the test treatment or the control is being administered. In addition, systematic methods have been introduced to ensure the quality of medicines; and in the field of biological preparations in particular, such as vaccines, Britain has set an example to the world as a whole in setting standards for effectiveness and safety. All of these developments have helped to demonstrate the medical efficacy of new treatments; but they throw no light on their financial costs or economic benefits.

Then in the 1950s, economists started to become interested in the subject of health care. They brought with them a fashionable economic tool called 'cost benefit analysis' (CBA). CBA provided a framework for measuring the overall economic benefits of medical care and setting them against the corresponding costs. These cost benefit studies were intended to be used to decide on the optimum allocation of resources between competing demands for government expenditure.

In practice, they turned out to be more in the form of public relations exercises to argue in favour of particular public projects, or to justify government expenditures *post hoc*.

In health, diseases such as tuberculosis provided splendid examples in which it could be shown that the cost of the modern anti-tubercular medicines was far outweighed by the savings in hospital costs and by the reduction in sickness absence and premature mortality. Such analyses yielded clear-cut evidence that the economic benefits from such treatments brought disproportionate savings in monetary terms. Just as clinical trials had demonstrated the therapeutic benefits of new medicines of the 1940s, early CBA studies demonstrated their economic benefits in the 1950s.

However, the early euphoria in which it appeared that modern medicine actually saved money was shortlived. It soon became clear that as modern medicines increased longevity and extended the scope of therapy, overall costs were going to rise out of all proportion to the calculable financial benefits in which they resulted. Thus in the 1960s the economic emphasis swung away from cost benefit analysis to the different concept of 'cost effectiveness analysis' (CEA).

To try to explain the difference, CBA attempted to set all the costs involved in a procedure against all its direct and indirect economic benefits. Hence – taking an example from a quite different type of activity from health care – a CBA study was able to show that the construction of the original motorways could bring economic benefits by speeding up the flow of traffic. Similarly, in medicine, CBA set out to answer the question 'Does this treatment bring a greater or a smaller economic benefit than its economic cost?' In the 1950s, it has already been pointed out that the answer was often 'yes'.

On the other hand, 'cost effective analysis' (CEA) attempts to answer the simpler question 'Which of these alternative treatments provides better value for money – given that they both achieve the same outcome?' Alternatively, CEA can tell health service planners how they can achieve the greatest benefit from a given quantum of resources. Thus it is not concerned with an overall balance between economic costs and economic benefits (like CBA) but is concerned with the most efficient use of resources given that either the desired outcome or the available amount of resources are already fixed.

CEA is still an essential economic tool in rational health care planning. However, it does not provide an answer to the broader question of how much overall benefit a service is yielding. In an attempt to tackle this question, in the 1970s economists started to apply another technique, of 'cost utility analysis'. Figure 1, therefore, shows a simplified history of the development of stages in the evaluation of health care between the 1940s and 1980s.

Figure 1 Development of economic measurements of health and the effectiveness of therapy.

<i>Year</i>	<i>Type of evaluation</i>	<i>The question it answers</i>
1940s/50s	Clinical trials	Does the treatment work?
1950s/60s	Cost Benefit Analysis (CBA)	Does the treatment pay off?
1960s/70s	Cost Effectiveness Analysis (CEA)	Which is the most effective treatment using given resources?
1970s/80s	Cost Utility Analysis (CUA)	How does the treatment affect the length and quality of life?

The concept of 'utility'

Cost utility analysis (CUA) introduces the economists' concept of 'utility' as a measurement of value to replace the financial measures embodied in the economists' classical concept of 'profit'. Very simply, the idea is that there is more to life than money. A person's quality of life may be more important than his material wealth.

Of course, the two can often go hand in hand. If a person can afford the money to install central heating, he will be warmer and more comfortable than someone who cannot afford it. Similarly, if people can afford to travel, their minds should be enriched. With elegant surroundings in their homes, their quality of life is also enhanced. But, on the other hand, a contented poor man may be 'better off' than an unhappy rich one.

There is nothing new in that philosophy, of course, but economists have attempted to quantify it by measuring the 'utility' of a person's existence – or a particular activity – independent of direct financial considerations. In other words, economists judge something's utility by how much it is wanted or valued, irrespective of how much may actually be paid for it. The relevance of this approach to health care in a welfare state is obvious.

As the historical outline above suggested, the principle of 'utility' has become increasingly important in the evaluation of health care since relatively few treatments now bring actual financial savings. In the 1940s many treatments reduced stays in hospital and eliminated costly premature deaths. Although this is sometimes still true in the 1980s, treatments are now more often concerned simply with making people feel better rather than with saving money. Clearly this 'wellbeing' has a utility value in economic terms, regardless of the fact the people do not have to pay directly to obtain it.

Thus cost utility analysis attempts to measure the degree of wellbeing achieved by a treatment in relation to its cost. And cost in this connection, under the National Health Service, of course, to a great extent, means the amount of tax-payers' money which is required to provide health care manpower and materials. In other words, CUA recognises that the health services do not exist to create wealth but exist instead to create wellbeing – or to improve the length and quality of life. Furthermore, unlike the early antibiotics and vaccines, many modern medical treatments are no longer primarily concerned with saving life or extending longevity. Their objective is to make patients feel better either during acute episodes of illness, or more often during the development of chronic progressive illhealth. Therefore, if economists are to make their maximum contribution to the planning of health services and to the rational allocation of health care resources, they now need realistic measures of wellbeing and quality of life as well as measures of survival and longevity.

Incidentally, these measures of the beneficial outcome of medical treatment are needed not only to justify its cost. They are also

increasingly needed to justify the risks inherent in medical treatment. Although medicines generally are becoming safer, much publicity has recently been given to the adverse reactions which have always sometimes occurred with the use of medicines. Too little attention has been paid in public discussion to the other side of the coin – the substantial benefits which medicines have achieved, both in saving lives and in making life more pleasant.

The techniques of measurement

There are three broad approaches which have been developed to measure the wellbeing (or lack of it) in patients. These are:

- a) disability scales
- b) health profiles
- c) health indices

The techniques as a whole are often described as the use of 'health indicators', and their historical background has been fully discussed in a recent book edited by Professor Culyer (1983) and in another published by OHE (Teeling Smith 1984). Both these books were based on the proceedings of conferences held on the subject.

Disability scales

One of the earliest attempts to scale disability was introduced in the 1940s by Guttman (1944). He ranked degrees of patient dependence in respect of a number of activities, such as feeding, continence, ambulation, dressing and bathing. Provided that dependence progresses steadily from one activity to another (ie, patients first have difficulty in bathing, then in bathing and dressing, and so on until finally they are dependent in respect of all five activities) this method of scaling disability would yield a single rating from one (no dependence) to six (dependent in all five activities). The resulting type of 'Guttman Scale' has been widely used since the 1940s (Culyer 1978).

Disability scales have also been developed for a range of individual diseases. One example is shown in Figure 2 (Kurtzke 1981). This has been prepared for use in multiple sclerosis; it lists twelve activities, giving a 0 to 4 rating for the patients' abilities for each activity. The patients' overall rating on the scale gives a measure of the extent of their incapacity as a result of the disease. It can compare one patient's condition with that of another; it can give a measure of the total extent of incapacity in a given community; and it can measure a patient's deterioration or improvement over a period of time, possibly as a result of appropriate medical intervention.

Other well-developed scales of disability have been produced for anxiety and depression. Figure 3 shows one example, which is the questionnaire which forms the basis of a Hospital Anxiety and

Figure 2 Incapacity scale.

Function	Performance				
	normal	Impaired but able to do:			
		with- out aid	with mech- anical aid	with humar aid	not at all
Score	0	1	2	3	4
1 Stair climbing					
2 Ambulation					
3 Chair/bed transfer					
4 Bowel function					
5 Bladder function					
6 Bathing					
7 Dressing					
8 Grooming					
9 Feeding					
10 Communication					
11 Physical problems (tentative)					
12 Social dependence					
Total					sum

Source Kurtzke (1981).

Depression Scale. This was developed in Leeds and is a self-assessment rating (Zigmond and Snaith, 1983). The form shown is available in a variety of languages. The answers are scored from 0 (best) to 3 (worst) and the total from each answer can be summed. This disability rating is, therefore, comparable to the health indices to be discussed later. Many other similar questionnaires have been produced for the assessment of effects of mental illnesses.

Although there is clearly an overlap between disability ratings produced for individual diseases and more general health questionnaires, the former usually have a more specific application. They may, however, yield an overall score, as in the three cases quoted above.

Health profiles

A more sophisticated and general approach to the measurement of wellbeing is embodied in the use of health profiles. These give quantitative measurements of wellbeing for a number of distinct

Figure 3 HAD scale.

Doctors are aware that emotions play an important part in most illnesses. If your doctor knows about these feelings he will be able to help you more.

This questionnaire is designed to help your doctor to know how you feel. Read each item and place a firm tick in the box opposite the reply which comes closest to how you have been feeling in the past week.

Don't take too long over your replies: your immediate reaction to each item will probably be more accurate than a long thought-out response.

Tick only one box in each section

I feel tense or 'wound up':

Most of the time
A lot of the time
Time to time, Occasionally
Not at all

I still enjoy the things I used to enjoy:

Definitely as much
Not quite so much
Only a little
Hardly at all

I get a sort of frightened feeling as if something awful is about to happen:

Very definitely and quite badly
Yes, but not too badly
A little, but it doesn't worry me
Not at all

I can laugh and see the funny side of things:

As much as I always could
Not quite so much now
Definitely not so much now
Not at all

Worrying thoughts go through my mind:

A great deal of the time
A lot of the time
From time to time, but not too often
Only occasionally

I feel cheerful:

Not at all
Not often
Sometimes
Most of the time

I can sit at ease and feel relaxed:

Definitely
Usually
Not often
Not at all

I feel as if I am slowed down:

Nearly all the time
Very often
Sometimes
Not at all

I get a sort of frightened feeling like 'butterflies' in the stomach:

Not at all
Occasionally
Quite often
Very often

I have lost interest in my appearance:

Definitely
I don't take so much care as I should
I may not take quite as much care
I take just as much care as ever

I feel restless as if I have to be on the move:

Very much indeed
Quite a lot
Not very much
Not at all

I look forward with enjoyment to things:

As much as ever I did
Rather less than I used to
Definitely less than I used to
Hardly at all

I get sudden feelings of panic:

Very often indeed
Quite often
Not very often
Not at all

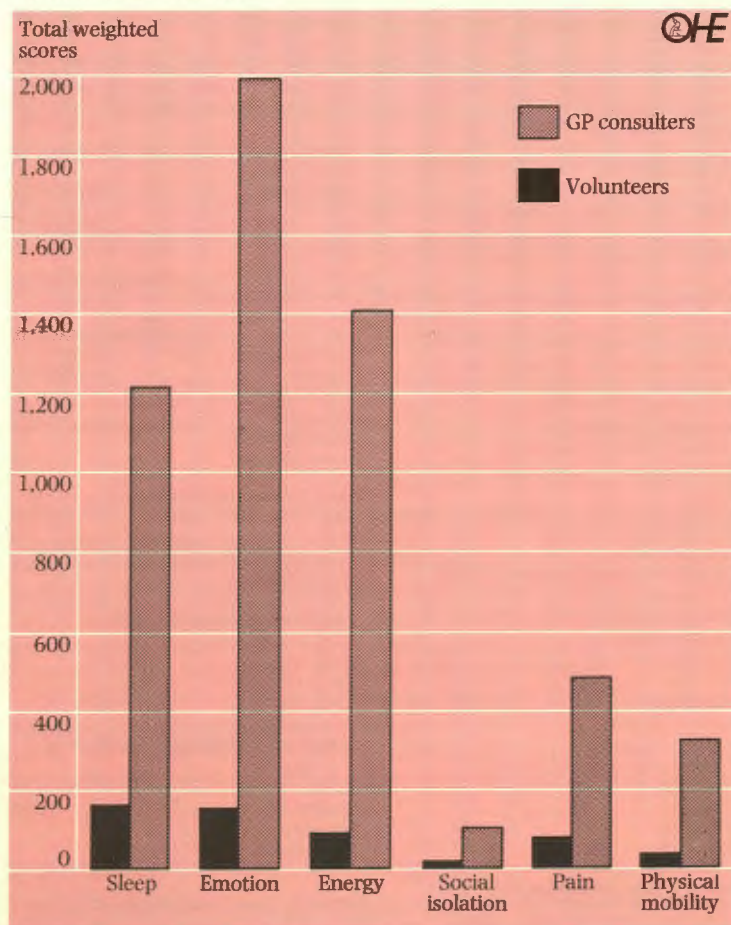
I can enjoy a good book or radio or TV programme:

Often
Sometimes
Not often
Very seldom

Source Zigmond and Snaith (1983).

variables, such as pain, tiredness, mobility and sleep patterns. The measurements for the different variables are not added together in any way, and hence in a real sense this approach produces a 'profile'. It may, for example, yield a high score for sleep disturbance, a moderate score for mobility, and a very low score for pain. The most important health profile to have been used and to have been extensively validated in Britain is known as 'the Nottingham Health Profile' (Hunt *et al* 1980). Figure 4 shows an example of the use of

Figure 4 NHP Pilot Study: comparison of responses between GP consultants and controls



this profile for patients consulting their general practitioners as compared to the profile for healthy volunteers (Stevens 1985).

This particular example uses the main section of the Nottingham Health Profile (NHP). This part of the NHP is based on the questions set out in Figure 5 under six main headings. Figure 4, therefore, shows the scores for each of the six main topics covered in this questionnaire. In addition, the profile includes seven 'yes/no' questions asking whether the following areas of daily life are affected:

- paid employment
- looking after the house
- social life
- home life
- sex life
- hobbies and interests
- holidays

In the main section the answers to the different questions set out in Figure 5 are given different weights according to their significance. Thus in the sections on physical mobility a positive answer to

Figure 5 Nottingham health profile: listing of statements.

Physical mobility:

- I find it hard to reach for things
- I find it hard to bend
- I have trouble getting up and down stairs or steps
- I find it hard to stand for long (eg, at the kitchen sink, waiting for a bus)
- I can only walk about indoors
- I find it hard to dress myself
- I need help to walk about outside (eg, a walking aid or someone to support me)
- I'm unable to walk at all

Pain:

- I'm in pain when going up and down stairs or steps
- I'm in pain when I'm standing
- I find it painful to change position
- I'm in pain when I'm sitting
- I'm in pain when I walk
- I have pain at night
- I have unbearable pain
- I'm in constant pain

Sleep:

- I'm waking up in the early hours of the morning
- It takes me a long time to get to sleep
- I sleep badly at night
- I take tablets to help me sleep
- I lie awake for most of the night

Energy:

- I soon run out of energy
- Everything is an effort
- I am tired all the time

Social isolation:

- I'm finding it hard to get on with people
- I'm finding it hard to make contact with people
- I feel there is nobody I am close to
- I feel lonely
- I feel I am a burden to people

Emotional reactions:

- The days seem to drag
- I'm feeling on edge
- I have forgotten what it is like to enjoy myself
- I lose my temper easily these days
- Things are getting me down
- I wake up feeling depressed
- Worry is keeping me awake at night
- I feel as if I'm losing control
- I feel that life is not worth living

the question 'I'm unable to walk at all' would clearly be given a heavier weighting than a positive answer to 'I find it hard to bend'. The profile shown in Figure 4 is constructed from the appropriate weighted answers to the questions under each of the six main headings. Figure 6 gives another example of the use of the main part of the profile, assessing the effect of pregnancy on wellbeing at 18, 27 and 37 weeks. Figure 7 shows the corresponding results from answers to the seven yes/no questions of the NHP (McEwen 1983).

Apart from the Nottingham Health Profile, other similar 'instruments' had earlier been developed in North America. One such health questionnaire was developed in Hamilton, Canada in the mid-1970s, and another at about the same time in Seattle (Rosser 1983). The latter was known as the 'Sickness Impact Profile' (SIP) and included both psychological and physical factors. The answers to its questions were grouped into fourteen categories, compared to the six in the NHP. However, as in the NHP, no attempt was made to aggregate the fourteen separate scores. Hence it, too, produced a 'profile' rather than a single 'index'.

Health indices

A more complex, and in many ways theoretically sounder, approach to the measurement and scaling of wellbeing has been the development of the general Health Index. These indices combine the scores

Figure 6 Median scores on Part I of the profile at 18, 27 and 37 weeks of pregnancy

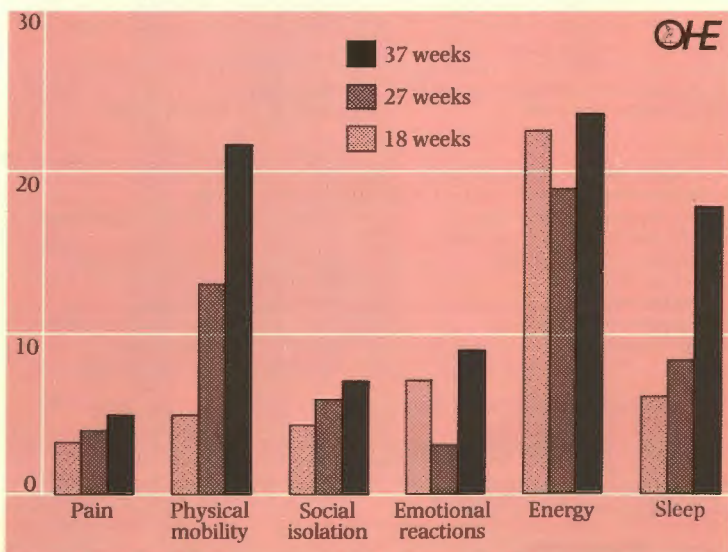
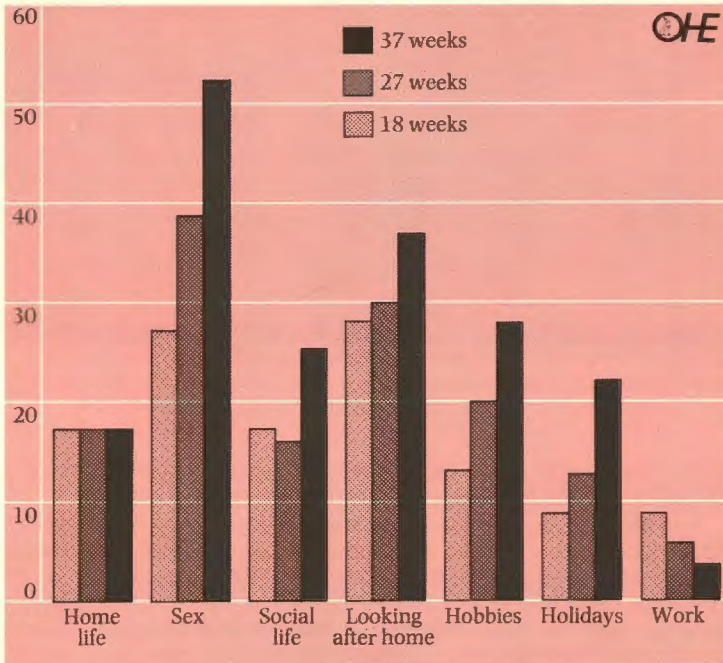


Figure 7 Percentage of women reporting problems in each of the seven areas of daily life at 18, 27 and 37 weeks gestation

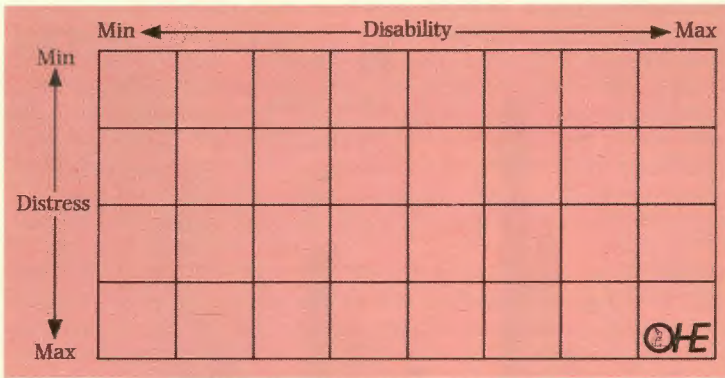


obtained for each of the measurements of disability and discomfort into a single global figure. Once again a number of different approaches have been adopted, but in each case the principle has been to take a value, for example One, to represent 'perfect well-being' and another value, for example Zero, to represent death. The range between these two points has been called the 'dead-healthy' scale.

In fact, one of the earlier pieces of work in this connection, by Fanshel and Bush in San Diego in 1970, extended the principle even further, and took into account the prognosis of the illness, as well as the patient's current state of health (Fanshel and Bush 1970). This, however, proved too complex, and subsequent developments have tended to concentrate only on the patient's present condition.

One of the earliest pieces of work in Britain was carried out by Rosser and Watts in 1971, at St Olave's Hospital in South London (Rosser and Watts 1974). They constructed a matrix showing four degrees of distress down one axis and eight degrees of disability down the other. They then allocated each patient to an appropriate

Figure 8 Matrix providing a 'sanitive index'



Source Rosser and Watts (1974).

box out of the 32 in the matrix (see Figure 8). The eight degrees of disability and the four degrees of distress are defined in Figure 9.

In order to put a 'value' on each box, they turned to a legal text book (Kemp *et al* 1967) which gives the quantum of damages awarded following industrial injury cases. They put each of about 500 legal cases into its appropriate box in the matrix, and derived an average value for the compensation given in each of the 32 different conditions. They converted these values into an index, and calculated a figure for the total 'wellbeing' in St Olave's Hospital by multiplying the number of hospital cases in each box by the appropriate index value for that box. They called the outcome of this exercise the 'sanative index' for the hospital.

They quickly recognised that this method of valuing different states of wellbeing was fairly naïve, and since the 1970s Rosser has developed a very sophisticated 'psycho-physical' method of valuing different states of wellbeing still using the same basic matrix (Rosser, 1984). Essentially this consists of describing examples of the various states, and using various devices to get different groups of individuals to put relative degrees of desirability on the different states. The process is described in more detail in a recent OHE research monograph based largely on the work of Torrance in Ontario (Holland 1985).

Still using the same matrix of disability and discomfort, Figure 10 shows the values attributed to each of the 32 states. It is significant that in two cases the score is less than zero; that is, the process of evaluation of the different states produced the conclusion that in these two states survival was worse than death. No scores are given for the three boxes in the matrix indicating a positive degree of distress in the eighth state of disability, as in this case the patient is

Figure 9 Rosser's classification of illness states.

<i>Disability</i>	<i>Distress</i>
I No disability	A No distress
II Slight social disability	B Mild
III Severe social disability and/or slight impairment of performance at work Able to do all housework except very heavy tasks.	C Moderate
IV Choice of work or performance at work very severely limited. Housewives and old people able to do light housework only but able to go out shopping	D Severe
V Unable to undertake any paid employment Unable to continue any education Old people confined to home except for escorted outings and short walks and unable to do shopping Housewives able only to perform a few single tasks	
VI Confined to chair or to wheelchair or able to move around in the house only with support from an assistant	
VII Confined to bed	
VIII Unconscious	

Source Kind, Rosser and Williams: 'Valuation of Quality of Life: Some Psychometric Evidence' in Jones-Lee M W (editor) *The Value of Life and Safety*, North Holland, 1982.

Figure 10 Psychometric scale of values of states of disability and distress.

<i>Disability</i>	<i>Distress</i>			
	1	2	3	4
1	1.000	0.995	0.990	0.967
2	0.990	0.986	0.973	0.932
3	0.980	0.972	0.956	0.912
4	0.964	0.956	0.942	0.870
5	0.946	0.935	0.900	0.700
6	0.875	0.845	0.680	0.000
7	0.677	0.564	0.000	-1.486
8	-1.028			

Source Rosser (1984).

unconscious. Hence by definition he cannot feel any conscious distress.

Various other health indices have also been developed, but the principle in each case is the same. A large number of different states of disability and discomfort are ranked on a one-dimensional scale, giving a relative value for each of the different conditions experienced by different patients – or by the same patient at different times.

The use of a one-dimensional scale provides for the possibility of adding scores for different individuals in a group, and – more importantly – for making a quantitative assessment of relative wellbeing for an individual over a period of time.

Measurements over time: the 'QALY'*

The assessments made using a health profile or a health index relate to a person's condition at a single point in time. They should exclude any consideration of future prognosis. Nevertheless, the measurements taken at different times can be used to provide a series of measurements over a given period. The consequent possibility of using a health index to provide measurements of the quality of life over a time span has led to the concept of a 'quality adjusted life year', or QALY. The principle is that it is no longer simply important to measure life expectancy in unadjusted years because a single year of excellent health may be equivalent, in terms of overall 'utility', to more than one year of impaired health. In other words, given the choice, a person would prefer a shorter healthier life to a longer period of survival in a state of severe discomfort and disability. The traditional emphasis on life expectancy has to be modified by 'adjusting' the number of years by a factor depending on their quality. A single year of better health may be equivalent in terms of its 'utility' to a longer period of survival with some pain, disability and suffering. In numerical terms, this means that a year and a half with a score of 0.66 QALYs per year would be equivalent to a single year of perfect wellbeing. Each would rate a score of 1.0 QALY, given that 1.0 QALY represents a year of unimpaired good health.

According to Figure 10 a state with a score of 0.68 would be represented by a person confined to a chair or wheelchair and in moderate pain. Eighteen months in that condition has the same economic 'utility' – on this basis – as one year completely free from disability or distress.

It must be clear that the extent of general acceptance of this concept depends on whether or not the scores given to different degrees of wellbeing can be agreed by different interested groups. Much work in Britain and North America is at present concerned with this question. There is, however, fairly general agreement that the basic principle is one which it is worth developing. A year of perfect wellbeing must have a greater 'utility' than a year of painful bed-bound survival for a seriously ill and suffering patient.

The practical problem of measuring different people's values for the same degree of disability immediately raises one of the problems involved in the practical use of health profiles and health indices. It may be difficult to get a universally accepted view of the value – or 'utility' – of different states of wellbeing. An elderly person may feel

their severely handicapped and painful existence is just as valuable to them as someone else's apparently healthier existence.

More generally, an obvious limitation of the use of measurements of health is that they discount other factors which affect a person's quality of life. For example, a sick man who is in a very satisfying job may be happier than a healthy man whose work bores him to distraction. However, this limitation is no reason not to try to develop valid measures of a person's state of health. It would be just as illogical to argue that one should not measure the quality of housing because it, too, was only one factor which affected a person's overall wellbeing.

At present the measurement of health in rigorous economic terms is in a state of development. Its usefulness is still only slowly being proved as a tool with which to demonstrate the relative value of different types of therapy. The general approach is, however, becoming increasingly accepted by doctors and economists working together. The next section of this report describes some of the studies which have recently been undertaken in this connection.

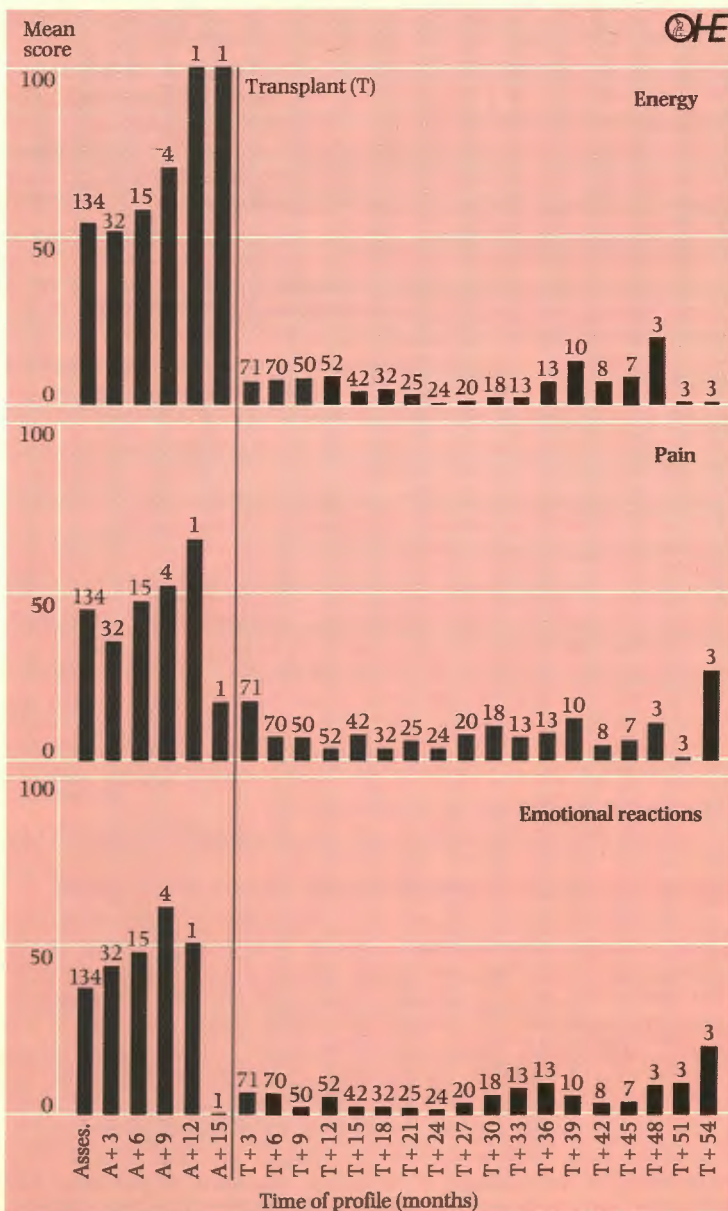
Practical applications

Figures 6 and 7 showed practical results from the use of the Nottingham Health Profile during pregnancy. As would be expected, the women were progressively affected as their pregnancy developed, and physical mobility was predictably most influenced. Sleep, also, was quite seriously affected, and the mothers-to-be suffered significantly from a lack of energy at all stages of the pregnancy. On the other hand, almost fifty per cent of women reported no problems with their sex life even at 37 weeks of pregnancy, probably because a negative answer would be appropriate if sexual activity was not taking place at this stage in the pregnancy.

Figure 4 showed the results of pilot studies recently carried out by the Centre for Medicines Research, where a research fellow is engaged full-time in developing the techniques of measurement.

A much more extensive use of the techniques for the measurement of quality of life have been used by SmithKline Beckman in the United States to evaluate the effect of their new oral gold preparation, auranofin, on patients with rheumatoid arthritis. These studies tested the sensitivity of a battery of Health Status measures in a 6 month multi-centre randomised clinical trial of auranofin versus placebo in 311 patients. Independent assessors, trained centrally, administered questionnaires selected to assess various dimensions of the disease. Four distinct dimensions were identified and represented by four composite scores. Changes in these scores were expressed as percentage of baseline standard deviations. The auranofin group improved significantly more than the placebo

Figure 11 Nottingham Health Profile, Part I: Mean scores for each section by three-month periods from assessment and from heart



transplant. Combined hospital data, (numbers of observations above bars)

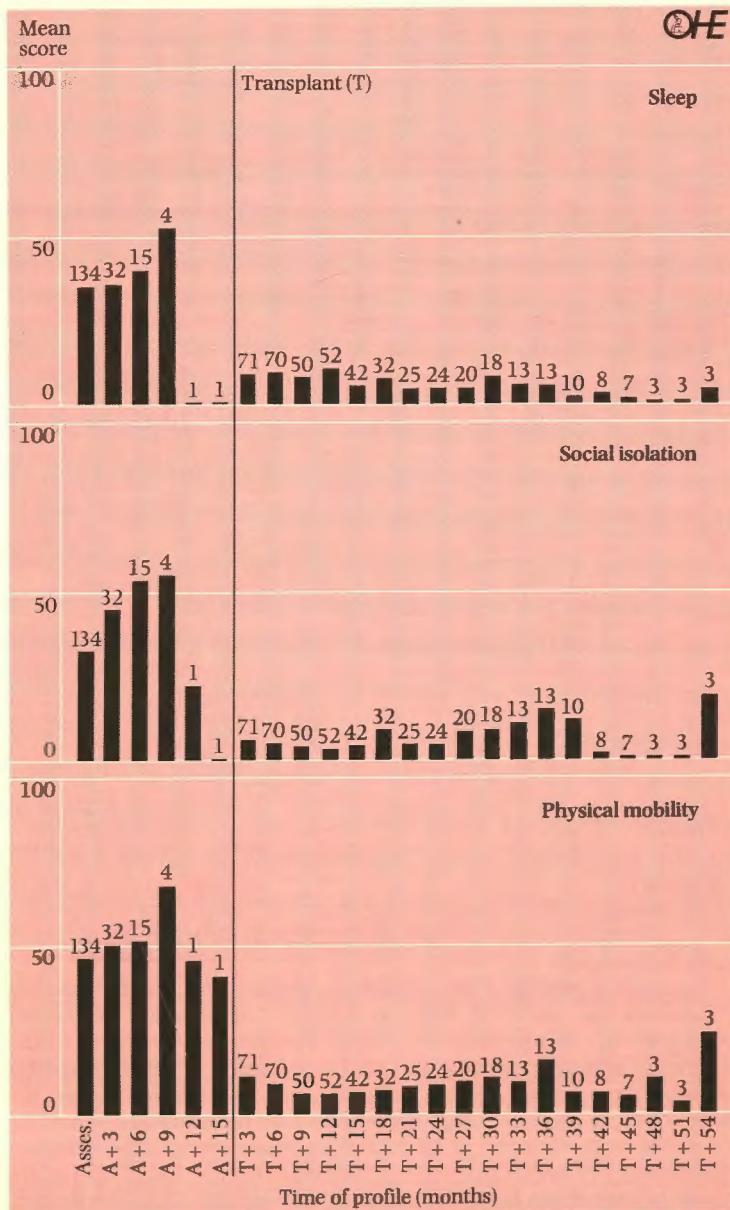


Table 2 Results of health status measurement on patients with rheumatoid arthritis.

<i>Dimension</i>	<i>Change in composite scores (%)</i>		<i>P value</i>
	<i>Auranofin</i>	<i>Placebo</i>	
Physical	34.5	15.7	0.003
Functional	29.2	4.9	0.001
Pain	74.1	49.7	0.026
Global	50.0	26.8	0.006

Source Auranofin co-operating group (1985).

group in all four dimensions. The results are shown in Table 2 (Auranofin co-operating group 1985).

A recent use of the Nottingham Health Profile has been the evaluation of the heart transplant programmes at Harefield and Papworth Hospitals by a team led by economists at Brunel University (Buxton *et al* 1985). The NHP was administered to all patients at their assessment, and then by post to accepted patients at three-month intervals while they waited for transplant. It was then similarly administered at three-monthly intervals post-transplant. Using all the available data on accepted patients the overall pattern of mean scores for each section gives a very vivid impression of the change that follows transplant. Figure 11 shows this for each of the six dimensions of Part I of the Profile.

Initially in setting up the study there had been some reservations about the sensitivity and accuracy of the NHP. In practice the results were found to tally well with other professional judgements and impressions. For example, the NHP data confirmed the distinction independently made by the surgeons at assessment at Papworth between definitely and provisionally accepted patients – the latter being patients who were judged to be in less immediate need of transplant. Mean rank scores for observations of these two groups of patients indeed showed that the provisionally accepted patients felt significantly less 'distress' in terms of all dimensions ($P < 0.01$). More generally, the various changes in NHP scores were consistent with the detailed material from in-depth semi-structured interviews with the patients.

In order to provide a benchmark for comparison, Buxton and his colleagues also used the NHP on a sample of coronary artery bypass graft (CABG) patients. For these patients too the effect of the operation was very positive: a paired comparison of before and after observations for 69 CABG patients showed that, although the improvement on the energy dimension was not statistically significant, the improvement in all other dimensions was ($P < 0.01$). A comparison between the pre-operative scores for heart transplant patients with those for CABG patients showed the CABG patients to

be significantly less 'ill' in terms of the three dimensions of energy, sleep and physical mobility ($P < 0.01$) and also social isolation ($P = 0.04$). There was no statistically significant difference in the other two. Post-operatively there was no statistically significant difference between the two groups on any of the six dimensions.

Theoretical implications

It was pointed out in the introductory paragraphs that the purpose of assessing outcomes of medical treatment was to decide on the comparative values of different health care activities, as well as to throw some light on the desirability of spending more or less on health care as a whole. It has been explained that the economist's technique for reaching relevant conclusions is the Cost Utility Analysis. This, in turn, has led to the concept of the Quality Adjusted Life Year as a measure of the reduction or increase in health-centred Utility, due respectively to disease and to treatment.

Cost utility analysis aims to measure the benefits on quality of life yielded by therapies. It is therefore necessary to construct an index which measures the demand for health care before and after therapy. The benefit yielded by the therapy is the difference between 'after' and 'before' on the index. That difference can be translated into a function of the known cost of the therapy and that function is the cost utility of the therapy. It is that function which brings subjective change in wellbeing into relation with the objective cost of that change.

At a recent meeting to try to reach a consensus on the policy to adopt towards coronary artery by-pass graft surgery, Professor Alan Williams used this research to present some fascinating comparative data for his assessment of the cost-utility measurements for various conditions (Williams 1984). Some of these conclusions are summarised in Table 3. It shows the current value of present and continuing treatment costs (if applicable) for nine different medical procedures. It also gives an estimate of the benefits achieved in each case. The value of future years benefit are appropriately discounted.† Professor Williams regards his figures as necessarily tentative. However, the difference in orders of magnitude for the cost per QALY must provide food for thought. Hip replacement and pacemakers seem to be the 'best buys' on this analysis, with CABG for

†The reason for 'discounting' future benefits is that – even if money keeps its constant value (without inflation) – a gain in future is worth less in real terms than a gain today. For example, a gift of £100 now is worth much more than the same gift of £100 ten years hence. This is because, if it is received now, it can be invested (or used in other ways) so that its value is increased as the years go by. If the £100 were invested, for instance, to obtain a compound interest rate of 10 per cent per annum, it would be worth about £260 in ten years time. Conversely a gift of £100 ten years hence would be worth only £38 today. Hence discounting is getting at the present value of a future gain.

Table 3 Summary of costs and benefits for various treatments.

<i>Treatment</i>	<i>Present value of service costs (£'000)</i>	<i>Discounted gain in QALYs</i>	<i>Cost per extra QALY</i>
Pacemaker implantation for atrioventricular heart block	3.5	5	0.7
Hip replacement	3	4	0.75
CABG for severe angina (with left main vessel disease)	2.85	3.5	0.81
Valve replacement for aortic stenosis	4.5	5	0.9
Kidney transplant (cadaver)	15	5	3
CABG for mild angina (with left main vessel disease)	3.15	1	3.15
Heart transplantation	23	4.5	5
Haemodialysis at home	66	6	11
Haemodialysis in hospital	70	5	14

Source Williams (1984).

severe angina close behind. On the other hand, renal dialysis appears to yield considerably less benefits in relation to cost than all forms of cardiac surgery, including heart transplants.

Obviously, these figures are open to challenge but they introduce an important new dimension into health care planning and the allocation of health care resources. Doctors and economists now need to work together to improve the methods of quantification in this sort of exercise, and then to stand back and take a long hard look at some of the conventional attitudes towards the value of different treatments. When the concept is further extended to the care of chronic disabling conditions, such as rheumatoid arthritis, a whole new area for debate on the allocation of resources is likely to be opened up.

In initiating this debate, it is important to understand the strengths and weaknesses of the concept of economic 'utility'. Economic 'utility' is not usefulness as judged by an objective standard but the *subjective* importance of goods and services to the consumer. Any measure of economic utility is therefore a measure of the consumer's *feeling* of desire for the goods or services – in this instance, health therapy – rather than a measure of objective benefit from them. The enjoyment of these benefits should have the effect of satisfying that particular desire. A successful course of health therapy should produce a degree of wellbeing that satisfies the desire for it, wholly or in part; that is, it should produce beneficial movement of the patient's health index.

Much economic activity is devoted to stimulating consumer desire for goods and services of all kinds and the growth in recent years of 'consumerism' demonstrates that only by adequate measurement

and informed foresight does the capacity to stimulate demand merge with the capacity to yield benefits in consumer satisfaction. This applies to health therapy as much as to other services. Systems of measurement which give 'scale' values of 'before and after' health therapy allow comparative assessments of the benefits of the therapies available. It follows that much better-informed discussions of health policies in patient care become possible.

It must be emphasised, however, that systems of health measurement aim at producing comparable facts and do *not* aim at proposing solutions to problems. Measurements give a data base for consideration but that data base needs to be considered in terms outside measurement – in terms, that is, of current medical ethics, current social attitudes and current political thinking. To these fields of thought measurement is neutral; a mile is a measurement; a 'mile too far' is a judgement outside the methodology of measurement. But the original measurement of the mile is a necessary prelude to the judgement. So, in health services, measurement by acceptable scales is a necessary prelude to rational judgement of modern therapies.

By the 21st century, it seems likely that health services will be being planned and organised with the concept of Cost Utility Analysis just as much an accepted and central feature as controlled clinical trials have become by the 1980s. Meantime, the philosophy underlying the relevant techniques and their practical applications should continue actively to be developed along the lines already established in Britain and in North America. The Office of Health Economics is making an important contribution in this connection through its support of studies of the subject in the Department of Economics at Brunel University.

In conclusion, it is important to emphasise again two points. First, there is still much distance to travel before generally agreed values for health indices are eventually developed. Different groups of individuals perceiving a medical problem from different viewpoints may award the same state of health different relative values. But given time and experience it is probable that generally accepted values will emerge in the measurement of health states. Professor Rachel Rosser, in the results of her work depicted in Figure 10, suggests that some states of survival have a negative score: that is, they are worse than death. Clearly more work is needed before the implications of that conclusion can be generally accepted. But it could, eventually, argue in favour of voluntary euthanasia.

The last point to re-emphasise is that the whole concept of health indicators – or the measurement of health status – can never provide more than a background to political and medical decision-making in health care. Just as some people choose not to purchase the cheapest shoes in the shoe shop, so health service employees and administrators may not always choose the 'cheapest' procedures as revealed by cost utility analysis. However, just as it is impossible to make a

rational decision about the best pair of shoes to buy without knowing their price, so it is impossible to decide on the optimum allocation of health care resources without having a measure of the 'utility' of different procedures in economic terms.

Scientific measurement is an essential precursor of the logical decision process. This report has attempted to describe the techniques by which more logical decisions can be reached through an awareness of the 'utility' of different health care procedures in relation to their cost.

References

- Auranofin Co-operating Group (1985). Sensitivity to Change of Health Status Measures in a clinical trial of auranofin in RA; *Arthritis and Rheumatism*; 28, 4, supplement 8-28.
- Buxton M J, Acheson R M, Caine N, Gibson S and O'Brien B J (1985). Costs and Benefits of the Heart Transplant Programme at Harefield and Papworth Hospitals. HMSO.
- Culyer A J (1978). Measuring Health; lessons for Ontario. University of Toronto Press.
- Culyer A J (Editor) (1983). Health Indicators. Martin Robertson.
- Fanshel S and Bush J W (1970). A Health Status Index and its Application to Health Service Outcomes. *Operations Research*, 18, 1021.
- Guttman L (1944). A basis for scaling qualitative data. *American Sociological Review*, 9.

- Holland G (1985). Techniques of Health Status Measurement Using A Health Index. Office of Health Economics.
- Hunt S M, McKenna S P, McEwen J, Backett E M, Williams J and Papp E (1980). A Quantitative Approach to Perceived Health Status: A Validation study. *Journal of Epidemiology and Community Health*, 34, 281.
- Kemp D A M, Kemp M S and Harvey R O (1967). The Quantum of Damages. Sweet and Maxwell.
- Kurtzke J F (1981). Initial proposal for a uniform minimal record of disability. *Acta Neurologica Scandinavia*, supplement 87, 64, 48.
- McEwen J (1984). The Nottingham Health Profile; a measure of perceived health. In: G Teeling Smith (Ed) Measuring the Social Benefits of Medicine. Office of Health Economics.
- Rosser R M (1983). Issues of Measurement in the Design of Health Indicators; a Review. In: A J Culyer (Ed) Health Indicators. Martin Robertson.
- Rosser R M (1984). A history of the development of health indicators. In: G Teeling Smith (Ed) Measuring the Social Benefits of Medicine. Office of Health Economics.
- Rosser R M and Watts V C (1974). The development of a classification of symptoms of sickness and its use to measure the output of a hospital. In: D Lees and S Shaw (Eds) Impairment, Disability and Handicap. Heinemann for the SSRC.
- Stevens J (1985). Private communication based on material published in *CMR News*. Centre for Medicines Research 1984, 2, 3.
- Teeling Smith G (Ed). Measuring the Social Benefits of Medicine. Office of Health Economics.
- Williams A (1984). Coronary Artery Bypass Grafting; an Economic Appraisal. Paper presented at the Consensus Development Conference on Coronary Artery Bypass Surgery.
- Zigmond A S and Snaith R P (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavia*, 67, 361.

OHE Publications

Studies of Current Health Problems ISSN 0473 8837

- 64 Scarce Resources in Health Care 60p
- 65 Schizophrenia: Biochemical impairments, social handicaps 60p
- 66 Dementia in Old Age 60p
- 67 Huntington's Chorea 60p
- 68 Leukaemia: towards control 60p
- 69 Suicide and Deliberate Self-harm 60p
- 70 Alcohol: reducing the harm 60p
- 71 Hip Replacement £1.00
- 72 Medicines, Health and the Poor World £1.50
- 73 Coronary Heart Disease £1.00
- 74 Pharmaceutical Innovation £1.00
- 75 Understanding the NHS in the 1980s £1.50
- 76 Childhood vaccination – current controversies £1.00

OHE Briefings

- Renal Dialysis 30p
- The Effects of Prescription Charges 30p
- Trends in European Health Spending 30p
- Disability in Britain – the process of transition 30p
- Sickness Absence: a review 30p
- Accidents in Childhood 30p
- Doctors, Nurses and Midwives in the NHS 50p
- Medicine and the Quality of Life 50p
- Ill in Europe 50p
- Keep on Taking the Tablets? 50p
- The Politics of Prescribing 50p

Reports on OHE Symposia

- Health Research in England: a topic for debate £2.00
- The Second Pharmacological Revolution £7.50
- Measuring the Social Benefits of Medicine £7.50
- A New NHS Act for 1996 £1.50

Pharmaceutical Monographs

- A Question of Balance: benefits and risks of medicines £1.50
- The Consumer Movement, Health and the Pharmaceutical Industry £2.00
- Issues in Development: a guide £2.00
- The Future for Pharmaceuticals £2.00
- Pharmaceuticals in Seven Nations £2.50

Compendia

- Compendium of Health Statistics. 5th Edition; 1984 £15.00

Research Monographs

Costs and Benefits of Regulating New Product Development in the
UK Pharmaceutical Industry £5.00

The Needs of the Pharmaceutical Manufacturers from their Medical
Departments in the 1900s £1.50

Patterns of European Diagnoses and Prescribing £5.00

Pharmaceuticals in Developing Countries; 1981-82 £2.00

Techniques of Health Status Measurement using a Health Index £2.50

About OHE *free*

Medicines: 50 years of progress *free*

