Resources and decisions in clinical radiology

THOMAS SHERWOOD

From the Department of Radiology, University of Cambridge

I am not an epidemiologist, a social physician, or an economist. My work is in clinical radiology: I x-ray sick people. I am, I suppose, representative of the everyday doctor who spends large sums of money on his patients each year.

The title given to me for this paper is not something that comes readily as part of a doctor's basic kit. Human interest, compassion, and curiosity are the sort of motives advanced by prospective medical students, but by long tradition these refer to the one-to-one relationship between the doctor and his sick patient. The Hippocratic oath says '... I will prescribe regimen for the good of my patients according to my ability ...'; For 2400 years doctors have cared for individual patients who seek help with an immediate problem. That may not look like a wholly admirable stance now, but it is an interaction which has been established between the individual in trouble and the doctor as an honourable professional caring for that particular man or woman. It is probably the way many of us would still like to be treated if we have the misfortune to fall sick.

It is of course easy to show that this traditional and respected picture is really riddled with holes. What compassion is there in caring extravagantly well for one patient, and then being unable to look after the next two? Or in tinkering with the late ill effects of disease instead of preventing it at source, like removing the handle of that cholera-infested pump in Broad Street? As a malicious aside, I want to remind community medicine men that even their best interventionist efforts, just like the clinician's, can come to look a trifle puny against the tide of natural history. It seems that the Broad Street cholera outbreak was already past its peak and on the wane when John Snow had that pump handle removed in 1854 (Bradford Hill, 1955).

To the clinical doctor with his professional ancestry of thousands of years, the about-turn from trying to relieve individual suffering to working for Health with a capital H is a revolution that has not come easily. Radiologists have been worried for longer than those in some other disciplines, because the gulf between resources and needs has been brought home to them on their own doorstep. Demands for x-ray services are growing at 5% to 10% compound interest each year. Two-thirds of the whole population in the United States of America are estimated to have had a radiological examination in 1970 (British Medical Journal, 1977). Forty million units of x-ray work were carried out in the National Health Service in 1974 (one chest radiograph is one unit). There are about 1000 pairs of radiological eyes at all grades in the NHS, and it is increasingly hard for them to cope with the flood of cases which face them. The gap between limited resources and open-ended demand, between what is economically feasible and medically possible, is very wide in radiology departments. No wonder that radiologists are interested in clinical decision-making, and have sometimes been in the forefront of exploring this process (Knowles, 1969; Lusted, 1971).

We cannot be endlessly generous and continue to be fair (Acheson, 1978). Health care has moved, never to return, along a historical path from charity, via market economy, to corporate provision. Traces of the earlier systems remain, but today we have to discuss how corporate provision in the NHS might be made to work. I would like to examine decision-making as a bridge between resources and demands by three radiological examples: the routine chest x-ray, the investigation of hypertension, and the problem of finding a lump in the kidney. These have been chosen because they illustrate attempts to come to grips with problems of audit, decision analysis, and, indeed, scientific method in clinical medicine.

THE ROUTINE CHEST X-RAY
ALL-OR-NONE TESTING

The taking of a chest radiograph is a hallowed ritual for every patient passing through the front doors of any hospital. When pulmonary tuberculosis was rife in the population there may well have been point in this sledge-hammer approach. Current conventional teaching still demands a chest x-ray as a routine baseline observation, so as to exclude any cardiothoracic disease there might be, and to monitor further progress. Follow-up visits, particularly to the cardiac or hypertension clinic,
tend to be an automatic signal for repeat examinations. Hospitals themselves often stipulate an entrance chest x-ray examination for their staffs, as certain governments do for immigrants. Good anaesthetic practice, it is said, demands a chest film as a preoperative measure for every patient.

The question to be asked here is initially in simple all-or-none form: does the routine chest x-ray generate a worthwhile pick-up of disease states? 'Worthwhile' is of course an arguable term—it might be used at first just for pointing to disease not discoverable by simpler routines, for example, by taking a history and examining the patient. Studies of this have been done. Sagel et al. (1974) in the United States of America, and in this country Kerr (1974) and Rees et al. (1976), have shown that the routine chest x-ray has nothing to offer the young adult in the absence of chest symptoms or signs. Even in older age groups, the significance of abnormal findings seen only on the chest x-ray is highly doubtful. In the Cardiff series (Rees et al., 1976) cardiac enlargement accounted for half of all abnormalities. 'So what?' one can reasonably ask. A large heart will be detected by clinical examination. History and examination together will sort out the usual important background to the finding: hypertension and/or myocardial ischaemia. While the 'worthwhile' aspect of this ritualistic procedure over the whole hospital age group remains debatable, therefore, certain helpful guidelines can be stated at once: (1) there is no point in routine chest radiography of young adults under 30 in the absence of symptoms or signs and (2) the pay-off for older adults is at best doubtful; certainly those under 40 do not need a second radiographic view (a lateral film) as part of the routine examination (Sagel et al., 1974).

**Radiology in Hypertension Decision Analysis**

In the late 1950s and early 1960s surgical and radiological advances encouraged the view that hypertension often had a surgically correctable cause: obstructive lesions in the main renal arteries. It was held that hypertensive adults under 40 without a family history of high blood pressure needed an aortogram to diagnose or rule out this possibility (McMichael, 1961). A critical and, for this period, courageous study showed that the cost/benefit ratio in terms of the patient himself was lopsided, with an appreciable hazard for very little clinical benefit available to those so investigated (Chamberlain and Gleeson, 1965). Even for young adults with high blood pressure, ostensibly the group which stood to gain most from intensive investigation, the clinical outcome was unimpressive (Breckenridge et al., 1967). Renal artery stenosis was found to be a frequent lesion in autopsy surveys, occurring in normotensive as well as hypertensive patients.

At least 10% of the adult population are hypertensive. They will need an intravenous urogram or a radioisotope renogram as a screening test if renovascular hypertension is to be discovered. A proportion of these tests will be found abnormal, but in this country this leads to a change in management of perhaps one per cent of patients (Lancet, 1975).

In this way very useful question marks can be raised over the value of investigative procedures, but their impact on the decisions to be made by a doctor facing the problem of his individual patient is unclear. Correctable renovascular hypertension as a cause-and-cure entity was a highly attractive breakthrough, supported by the undoubted benefit gained by individual patients. Although two decades later the romance is over and the cold wind of realism is blowing, it does not readily tell us what to do. How do the individual patient's life chances compare if high blood pressure is expensively investigated first, or if it is treated blindly by drugs?

Looking at the intravenous urogram and radioisotope renogram in patients with high blood pressure, we are faced by expensive, fallible pointers to renovascular disease, which may or may not be present. Even if the disease is found, the consequent hypertension may or may not be surgically curable. McNeil and her colleagues at Harvard have done very interesting work on the outcome in human and cash costs of these investigative and therapeutic decisions (McNeil and Adelstein, 1975; McNeil et al., 1975b). I will not rehearse their papers in detail, but will refer to some of the basic tools used in such analyses (McNeil et al., 1975a). These include the ideas of sensitivity and specificity of diagnostic tests. Sensitivity expresses the true positive ratio of the investigation, measuring the fraction of patients with the disease who will be detected by an abnormal test result. Specificity expresses the true negative ratio of the investigation, that is, the fraction of patients without the disease who will be identified by a normal test result.

The tests in question, like most others in clinical medicine, do not produce black or white binary answers. They produce a range, and cut-off points can be chosen at will. If we face a treatable, potentially disastrous disease, we shall want to choose cut-offs at high sensitivity, in other words we want to pick up as many patients with the disease as possible, even at the cost of including a
number of those without it (poor specificity). For less serious diseases or more dangerous treatments, we shall be particularly concerned with specificity, perhaps accepting poor sensitivity, which will mean missing some patients with the disease. A receiver operating characteristic curve can be plotted between true positive and false positive test ratios, allowing examination of such cut-off points.

More complex parts of information theory can be used to calculate post-test probabilities about the absence or presence of the disease in a particular patient, once a diagnostic test has been characterised in this way. It is clear that these ideas presuppose a knowledge of the true prevalence of the disease in the population to be tested. We can then make useful statements. For example, if we use the urogram for finding renovascular disease and set sensitivity to detect 78% of patients with the condition (acceptable), this also means that an equal number of patients without the disease have an abnormal urogram, and are selected for pointless arteriography (unacceptable). It is not that the urogram is bad, but that most hypertensives do not have renovascular disease.

I have dwelt at some length on these methods, to show that tools are available for tackling the diagnostic worth of complex clinical tests. To complete the picture in hypertension, we have to consider the outcome of surgery in renovascular disease and of ‘blind’ medical treatment of all hypertensive patients, irrespective of the diagnosis. We also want to know the cost of diagnosis and of the various forms of treatment. What does all this add up to?

Provisos are needed. We shall exclude the rare metabolic causes of hypertension, and also coarctation. Children with high blood pressure are outside this discussion. Hypertension wholly resistant to good medical treatment will always require second thoughts. Within these limits, the Harvard work has helped to support the informed consensus that the middle-aged and older hypertensive patient is best treated medically in the first place, regardless of the underlying disease. This consensus will take care of the majority of patients, with 40 years as the arbitrary harbinger of middle age. It may come to apply to all adult hypertensive subjects. The consensus is not determined simply by the 1974 dollar cost of finding ($20000) and surgically curing ($200000) each patient with renovascular hypertension. Nor is it just the forbidding economic burden of screening and surgically treating all hypertensive patients with renovascular disease, costed at $10 to $13 billion for the United States of America. For the doctor facing his individual hypertensive patient, there is the realisation, derived from such studies, that life chances depend much more on age, sex, diastolic blood pressure at presentation, and compliance with medical treatment than on the hunt for surgically correctable lesions. There is therefore no point in carrying out urograms or renograms in adult patients presenting with high blood pressure; it is an unrewarding practice for the individual and for the NHS.

It is idle to pretend that this consensus is widely recognised, even in well-informed centres (see ‘Audit and pitfalls’ below). X-ray departments still receive requests for intravenous urograms in patients aged 60 with hypertension. This is because the intravenous urogram is part of the ‘diagnostic work-up’ for hypertension. I would like to examine this idea in the next section on renal masses.

**Finding a lump in the kidney**

**Scientific method**

A renal lump is now most commonly a chance finding during an intravenous urogram carried out for some quite different reason, perhaps because of bladder outflow obstruction in a man. The mass could be an early carcinoma, as yet symptom free, potentially curable, and well worth accurate diagnosis. It is more likely to be an unimportant, simple cyst. Other less common diagnoses are possible, for example an abscess, or haematoma. The conventional approach to this vital differentiation is the diagnostic work-up, meaning that we shall gather all the information we can about the lesion, weigh the evidence, and come to a conclusion about what can best be done for our patient. So we will do an ultrasound scan, radioisotope scan, renal arteriogram and venogram, lymphogram, and also a computed tomogram (or EMI) scan. These investigations all have their particular advantages and disadvantages in the diagnostic process, and some of them are uncomfortable and carry risks. Such a programme is therefore expensive (it costs perhaps £300 to £350) and certainly unkind to the patient. It also appears somewhat uncritical to the doctor. But it has an air of scientific respectability, because is not the impartial gathering of evidence followed by its interpretation the very stuff of scientific method?

It is not, of course, either in logic or in practice. Popper’s view of science (1959) as a trial-and-error activity has made considerable headway in the past 40 years. Science begins with a throw of ideas, and critical experiment and observation follow. It is a practical view (Medawar, 1969), with important consequences for medical learning and teaching, particularly in radiology (Sherwood, 1978).
The alternative attack on the problem is to build on the hypothesis of 'the next likely diagnosis', forcing a critical diagnostic pathway through the investigative jungle. We shall do ultrasound scanning, which is a simple outpatient test, and divide the patients on this basis into those with a probable simple cyst, and those with more complex lesions. Only the minority with complex lesions will need expensive inpatient investigation by an arteriogram to look at the next likely diagnosis: a renal tumour. If the arteriogram supports this idea, the path to surgery is clear. If it does not, the next most likely diagnosis on the arteriogram findings will be pursued by an appropriate investigation. The ultrasound diagnosis of renal cyst, made in the majority of patients, will require support from percutaneous puncture of the lesion. This aspiration of cyst fluid will again be done as an outpatient procedure under local anaesthesia. No further measures are needed if the diagnosis is confirmed in this way.

Such a critical diagnostic pathway is a trial-and-error exercise, with provisions for the mistakes which are bound to occur. False ultrasound diagnoses will be made. The patient with a renal cyst falsely read as a complex lesion on ultrasound will have an unnecessary arteriogram revealing the true diagnosis—a pity but not a disaster. More worrying is the patient with a renal carcinoma falsely read as a cyst on ultrasound. This patient will have his tumour needled, revealing at once that all is not well. Careful follow-up has shown that the patient's chances are not impaired by fine needle diagnosis of his renal cancer. Indeed, any scheme such as I have outlined demands that the experiment is then made of testing the hypothesis. This has been done now in several centres, showing a pleasing accuracy and safety to this particular pathway (Sherwood, 1975).

The justifications for constructing critical diagnostic pathways are kindness to the patient and scientific method. The diagnostic work-up is not only unkind and uncritical, it is unscientific—much the worst indictment. Radiologists wielding scarce, expensive, resources are at present beset by the diagnostic work-up as a major source of waste and misdirection in clinical decision-making. Each successful diagnostic advance naturally moves from yesterday's research procedure to today's routine investigation. The standard paper reporting it will say 'This important method, while not replacing existing tests, adds vital new information'. Anxious to do the best for our patients by scientific medicine, we rightly aim to be up to date. So we tend to end our case histories with an impressive diagnostic work-up for excluding all the conditions the patient might conceivably have, but probably does not have. Indeed, the chief on his ward round may upbraid his junior staff for leaving gaps in the diagnostic scenario, for example, for not requesting a barium enema for a patient with rectal bleeding who has not yet had a sigmoidoscopy.

I would like to explain that a costly recent addition to the radiological armamentarium, computed tomography or the EMI head scanner, has shown considerable promise in replacing several existing but unpleasant, expensive, and less informative diagnostic tests (Thomson, 1977). But if very many more patients are subjected to the next test, the cost advantage will disappear. Care- and cost-effective studies on the whole body scanner are awaited with interest (Abrams and McNeil, 1978).

A major redirection of teaching will be needed in medical schools to discredit the diagnostic work-up. Established critical diagnostic pathways building on 'the next likely diagnosis' are as yet few and far between. We may note that in such pathways, a normal finding need not be a slap in the face of the doctor, implying that he has guessed badly, or is playing the exclusion game. It can be used as an incisive tool for recasting probabilities as to the next likely diagnosis (Gorry et al., 1978). From every point of view, not least that of resources, we are better off looking for these pathways in teaching and research than continuing with the present tired practices.

MEDICO-LEGAL ASPECTS
The fear of leaving something undone which might lead to legal action hangs most heavily over medical decisions in the United States of America but is now often used to justify an obligatory flood of investigations in this country too. 'We must take a skull x-ray of every head injury, just in case . . .—yet knowing that the impact on clinical management from such investigations is nil except in sharply defined clinical circumstances, depressed fracture, for example. A coroner, Pilling, has pointed out that we have here made a rod for our own backs. The law is concerned with standards of reasonable medical care, and these are set by the profession. If x-rays or other investigations are requested as a matter of routine rather than for a reasoned clinical decision, we are skidding down a slippery path, squandering resources where they are not needed. 'The law will ask the right questions if the criteria for radiography are changed . . . If it becomes impossible for an expert witness to say that the normal practice would have been to request radiology, and he becomes obliged to say that the clinical findings would dictate the course of action,
Auditing and Pitfalls

I have touched on one or two corners of making decisions in a hospital-based diagnostic discipline within the NHS. The money saved, the resources freed for better things from illuminating these corners, would not amount to very much at first. A contributor to this issue has written challenging ‘Rational allocation of resources requires a decision theory model, which in turn demands some monetary valuation of human life’ (Card and Mooney, 1977). My canvas is smaller. This is not just because I wanted to write about what I know at first hand, but because I am a little wary of tackling problems by means of grand schemes. I was interested to visit the department of one of the authors cited earlier who has written excellently on the subject of better decision-making in radiology. Wandering round with a registrar before seeing the chief, I found patients having x-ray examinations under the very clinical circumstances condemned by this author. Confronted, he/she said ‘Well, I don’t meet all the clinicians here, and I suppose they haven’t read my papers’. At Cornell University a departmental utilisation review and medical audit has been in progress, and has itself been analysed (Lancet, 1976). Outside influences on this project first led to an increase of 20 in the non-medical staff and from three to 26 pages for the guideline documents. Length of hospital stay was one part of the audit: six patients were identified who had stayed too long. The cost of identifying each was about £34 000. It may be unfair to pick on such an isolated though well-studied example of the folly of unwieldy centralised audit, but my sympathies are certainly with Black and Rivers (1977) ‘... the best road ahead probably lies in building on our existing system of allocation of resources, accepting and testing, as they become available, refinements in decision theory and economic analysis’.

Monitoring one’s own performance appears an absolutely essential part of understanding how decision-making affects available resources. Even in a hospital department such vital feedback is hard to come by, and I understand the frustration of central planners who lack adequate returns on which to base decisions. My persuasion would be that local schemes of examining performance at a mundane, everyday sort of level are a good starting point for effecting change. In some centres they have existed for a long time as ‘death meetings’, reviewing complications and possible errors of judgement. Irving and Temple (1976) have discussed how such local audits can be run at district hospital level, and they have outlined the ground rules. The proof that such small-scale exercises on decision-making and resources are successful will be slow in coming. However, I do not think it likely that improvement can even begin to arise without them.

I want to highlight one small experiment carried out in my own region. In a district hospital there, the two radiologists have looked at their personal resources in terms of radiological manpower as a starting point. They have talked to the consultants and general practitioners who refer cases from inside and outside the hospital, telling them that so many thousand units of radiological work are available each month. It has been agreed that first call on these resources goes to all urgent requests, which are carried out at once. What is left is spread out among the outpatient appointments. The requests from each consultant’s team and from each general practitioner are monitored, and could be made available if needed. The clinical teams know that if they are profligate with requests for urgent investigations, their outpatient waiting list will lengthen. Any individual who might be tempted to bypass the agreed system by sharp practice knows that his use of the x-ray department through the inpatient or outpatient door is on record. The scheme has not been going for very long and could be criticised on various points, but it is a united local effort to make sensible decisions on radiological resources.

Conclusion

I have described a few everyday diagnostic problems in the NHS. There are very many more. The examples try to illustrate sensible and fair ways of tackling questions like ‘What is wrong with my patient, how do I find out, and what is the use of that anyway?’ Some look upon a what-is-wrong approach as running counter to the whole idea of health care, smacking of repair work on the human frame when we should be bent on building sound lives in splendid biological harmony. Short of these Utopias of health, all is not lost. Cheerfulness will keep breaking through, or perhaps my examples were just chosen with a large unconscious bias. It turns out that for each of those examples the diagnostic pathway that is most sensitive to the husbanding of resources is also the best one on
grounds of humanity and scientific method. Perhaps all we really need to learn and teach is good medical practice. Judgement is a decisive ingredient of good medical practice and Hippocrates had it right after all, because I quoted the extract from his oath incompletely at the beginning of this paper. The full sentence runs: ‘... I will prescribe regimen for my patients according to my ability and my judgement and never do harm to anyone’.

Comment: JEREMY HURST
Economist
Department of Health and Social Security

Few will wish to disagree with Sherwood that radiology presents us with stark economic problems. Many more patients than now, not only those with disease but those who can be told they are without it, might benefit from the speed and accuracy of radiology in clinical diagnosis. Recently, new techniques such as CAT scanning and ultrasound have added much to these potential benefits. Existing benefits, however, are bought at a high price: some £56m of recurrent expenditure and over £20m of capital expenditure in England alone in 1976–77. Also, radiology carries risks. Medical radiation is, for example, the principal man-made source of radiation exposure. Although the volume of radiology tests has been growing at more than five per cent per annum for some years, resources remain rationed. In terms of radiologists per million inhabitants, for example, Britain seems to occupy a lowly place among developed countries in the international league table. Despite this, radiologists have been to the fore in arguing that several radiology tests are unnecessary (Bull, 1974).

Are we making the right allocation of resources to and within radiology? The waste of routine chest x-rays on symptomless individuals will be familiar to health economists who have read J. D. Pole’s classic paper on the decision to run down the mass radiography service (Pole, 1971). This showed that, on the most favourable assumptions about benefits, the ratio of benefits to costs of mass radiography by the end of the 1960s was, at most, 0.5. The use of radiology to detect renovascular hypertension is based upon a sophisticated literature developed by radiologists with which some health economists may be unfamiliar. This literature uses statistical decision theory, information theory, and cost-effectiveness analysis to compare the value of particular tests for particular conditions with their costs. Health economists will feel at home with the tools of analysis deployed here, if not with their specific applications. In his last example, investigating a lump in the kidney, Sherwood draws upon the philosophy of Karl Popper to argue that the traditional ‘diagnostic work-up’ is inductive and unscientific. British health economists are probably Popperians to a man, but some may need persuading that the blunderbuss approach to diagnostic testing cannot be reconciled with the spirit of Conjectures and Refutations (Popper, 1963). What if tests are cheap and harmless and the time of patients and staff is dear? May not the blunderbuss (or pathology autoanalysers) be acceptable in terms both of scientific method and economics?

Sherwood eschews the use of grand schemes in tackling the economic problems of radiology. Rather, he suggests that we adjust the existing allocation of resources gradually, placing emphasis on local clinical audit schemes. There is much to recommend such a judicious approach. The behaviour of clinicians is one of the main keys to understanding the use of tests. A series of recent American studies, admittedly in one hospital only, showed that the use of laboratory tests, including x-rays, varied as much as twentyfold between physicians. (Schroeder et al., 1973; 1974; Schroeder and Daniels, 1977). When the cost of tests was compared with rough measures of the quality and outcome of care between physicians, no significant correlation was found. In one of the studies, the physicians providing high quality care tended to cluster around the mean in their use of tests. The physicians providing low quality care tended to go to extremes. It would be interesting to know whether such findings could be reproduced in Britain.

Local clinical audit may be a necessary condition for better use of radiology resources but is it sufficient? Consideration might also be given to improving information and incentives.

Sherwood argues the case for improving information about critical diagnostic pathways and the cost-effectiveness of tests. Perhaps we should be spending more on this type of evaluation research, even at the expense of a small part of the annual growth in radiology. Without such research, the local clinical teams may be adrift in a sea of uncertainty about benefits, costs, and probabilities. The point will be well taken, of course, that research is of little value if the findings are not taken up by clinicians.

Are there adequate incentives, however, for clinicians to take account of economic factors in their use of radiology tests? At present the making of requests may require little more than the placing of ticks on a form. If there is some chance, however small, that a test will benefit a patient and the cost is apparently negligible, is it any wonder that there is an excess demand for tests in many radiology departments? One possible solution to this problem is the use of clinical budgets (Wickings, 1978). Here, clinical teams would be given nominal budgets for the ‘purchase’ of diagnostic tests and other resources, together with information about their costs. Any underspending (or some agreed proportion of it) might be devoted to improvements in clinical services. In this way, the true repercussions of ordering ‘unnecessary’ tests would be made plain in the form of reduced services. If arrangements such as these became acceptable to clinicians, they might go some way towards resolving the dilemmas that Sherwood has placed before us.

Reprints from Professor Thomas Sherwood, Department of Radiology, University of Cambridge, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 2QQ.
References


