An Approach to Screening Policies

THOMAS McKEOWN, MD, FRCP

Professor of Social Medicine, University of Birmingham

THE NEED FOR A POLICY

The question inevitably asked about screening policies is whether they are necessary. Those who put this question usually see the need for control of some medical procedures: for example, when serious risks may be involved, as when immunisation against poliomyelitis was introduced; or when material is in short supply, as when antibiotics first became available; or where assessment of the balance of advantages and disadvantages is complex, as it clearly is in routine vaccination against smallpox at the present time and against tuberculosis in the foreseeable future. But it is by no means obvious to everyone that a national approach is required for screening, which many people think of as an extension of clinical practice. If they see no grounds for policies in relation to clinical procedures in general, they are unlikely to recognise that they are needed in screening. From this viewpoint it is sufficient if the clinician is convinced from personal experience that investigation is worthwhile, and the responsibility of the National Health Service is not to assess and advise on procedures, but merely to provide the resources required for their development.

In this article I shall not discuss the notoriously sensitive question of whether some control is desirable in respect of clinical procedures in general; whatever the decision on this larger issue, there is little doubt that a national policy is needed on screening. The grounds for this conclusion are that screening is distinguished sharply from ordinary clinical practice, a distinction obscured by the way in which it is usually defined.

Most definitions of screening are in terms of its aims; it is commonly defined as identification of disease that is unrecognised or in the presymptomatic stage. While such definitions are quite accurate, they do not focus attention on the feature of screening that distinguishes it clearly from other forms of medical practice. The distinction turns, not on its aims, but on its sponsorship, and arises essentially from the difference in the position of the doctor in relation to the patient in the two cases: in the one where the patient seeks the assistance of the doctor; in the other where the doctor undertakes to identify the patient who needs his assistance.

When the patient seeks medical advice, the doctor's ethical position is relatively simple: he attempts to do his best with the knowledge and resources available to him. He cannot fairly be criticised when the state of medical knowledge does not
enable him to diagnose accurately or treat effectively the condition about which he is consulted; nor can he undertake in all cases to provide the best skill and the full range of facilities from which his patient might conceivably benefit.

The position is quite different in screening, when a doctor or public authority takes the initiative in investigating the possibility of illness or disability in people who have not reported abnormal signs or symptoms: there is a presumptive undertaking, not merely that abnormality will be identified if it is present, but that those affected will derive benefit from subsequent treatment or care. This commitment is at least implicit, and except for research or the protection of public health, no one should be expected to submit to the inconvenience of investigation or the anxieties of case finding without the prospect of medical benefit. This obligation exists even when the patient asks to be screened or to have a health examination, for his request is based on the belief that the procedure is valuable, and if it is not, it is for medical people to make this known. The ethical considerations are so fundamental to the concept of screening that it seems essential to reflect them in its definition, as suggested in the review, sponsored by the Nuffield Provincial Hospitals Trust (1968): medical investigation that does not arise from a patient’s request for advice about specific complaints. This was considered to include patients who have not sought medical assistance (as with mass radiography), who have sought medical assistance only for a screening test of health examination which they believe will be of value to them, and who have sought assistance for a condition unrelated to a screening procedure. This interpretation regards as the unique feature of screening the fact that investigation is initiated by the doctor rather than by the patient. It is ‘the medical initiative which creates the obligation which . . . makes a strict validation procedure essential’.

There is another consideration that makes it imperative to adopt for screening a more rigorous approach than that which is followed in ordinary clinical practice: this is the scale on which a programme may be, and, if it is properly validated should be, applied. The application of a new therapeutic measure is developed gradually because its use is determined by the practising doctor who assesses its merits in relation to the alternative treatments available to him. But a screening programme believed to be of value invites large-scale application that may involve deployment of extensive resources at regional and national levels. Inevitably it is developed at the expense of other possible uses of health manpower and equipment; for example, a decision to introduce national screening for breast cancer would make heavy demands that could be met only by substantial new expenditure or by diverting staff and facilities from other work. This diversion may be justified, but it is clearly essential to show that it is before accepting breast cancer screening as suitable for widespread application. The main tasks confronting those concerned with the practice of screening are, therefore, to decide on an appropriate validation procedure, and to apply it to existing programmes and to new programmes before they are brought into general use.
EVALUATION OF SCREENING
The Nuffield review of screening outlined an evaluation scheme with the following requirements—

Definition of the problem. Screening programmes are sometimes difficult to evaluate because they have no clearly defined aims. It seems essential to remove ambiguity by clarifying at the outset the abnormality to be detected, the treatment to be offered, the groups to be screened, the tests to be used, and the stage of the disease at which they are aimed.

Review of the position before screening. To evaluate the contribution of screening it is necessary to assess the position before it is introduced, particularly knowledge of the disease (its prevalence, natural history, etc.) and the effectiveness of existing preventive and therapeutic measures.

Review of the screening procedure. There should be an appraisal of the diagnostic methods to be used in screening (their error rates, applicability to the population under investigation, cost, etc.) and of the treatment or care it is proposed to offer to those identified as abnormal.

Summary of conclusions concerning the screening procedure. It is desirable to prepare a balance sheet, assessing the contribution of screening in relation to its costs, risks, and displacement of other medical measures.

Proposals for organisation of further evidence and initial applications.

The review outlined above summarises the current position and reveals areas in which knowledge is seriously deficient. With this information it should be possible to make proposals for additional investigations and to recommend the policy to be adopted in relation to the screening procedure. Broadly this might be (1) that the procedure is ready for general application, (2) that limited application is needed for further study of medical and/or managerial problems, or (3) that present evidence does not justify the introduction of a screening programme, and additional research is required.

In the Nuffield exercise an evaluation of this type was applied to a number of screening programmes in current use; several could not be justified on present evidence. Employed in this way, the procedure exposes deficiencies of knowledge and focuses attention on major requirements that should be met before a screening programme is accepted for general use.

However, while this approach is valuable, it is arguable that it is too exacting to be applied rigorously in all cases. In the history of medicine, measures have often been effective before they were understood: for example, improvements in hygiene were advocated and partially implemented without knowledge of the nature of infectious disease; withdrawal of thalidomide prevented malformations before the mechanism of teratogenesis was clear; and mortality declined in British doctors who stopped smoking in advance of a full picture of the aetiology of lung cancer. Not unreasonably, it may be suggested that if early detection of breast
cancer were shown to reduce mortality, the introduction of screening would be justified without insisting on a complete understanding of the disease, or even a critical appraisal of the results of treatment. The point is important because extensive and prolonged investigation is usually needed to unravel the natural history of disease, and there may be ethical and other objections to an objective assessment of treatment.

However, while we may accept that we do not always need to know how measures work, in view of the ethical obligations of screening we do need to know that they do work. We must also be sure that they do not expose patients to unknown or excessive risks, and that a new service will not be introduced at the expense of alternative and more effective uses of the same resources.

A NATIONAL APPROACH
An approach of the kind outlined above has been followed in Britain in planning screening policies since 1968. In that year the Joint Standing Sub-Committee on Screening in Medical Care was established by the Standing Medical Advisory Committees of England, Wales and Scotland: to review the field of diagnostic screening of the population for disease, to identify areas of needed research (co-operating where indicated with other departmental advisory committees), to consider implications for resources; and to advise the Standing Medical Advisory Committees of the Secretary of State for Social Services and the Secretary of State for Scotland on the justification for and operation of screening services. Its members have varied backgrounds—administrative, clinical, epidemiological—but have in common their interest and experience in the problems of screening. The Committee is strongly supported by government departments, and has observers from the Medical Research Council and other organisations concerned with screening.

The role of the Committee is essentially advisory. Questions related to screening are referred to it by the Departments of Health, but its main role has been to examine procedures and to advise on policy related to them. This advice normally takes the form of a report to the Standing Medical Advisory Committees in which a screening procedure is evaluated and recommendations are made in relation to service policies and requirements for further research. After review by the Standing Medical Advisory Committees the implementation of the recommendations are matters for the Departments of Health and (in respect of research) the Medical Research Council.

Since the resources of the Committee are limited—it has no full-time staff—it must rely on the efforts of its members and others, and on the support given by the Departments of Health. When a screening procedure is considered to require evaluation, the normal practice is to establish a small working group composed of members of the Committee and of experts on the subject who are not members. The group prepares a report which is considered by the Committee and (usually) revised before submission to the Standing Medical Advisory Committees.
In the past seven years many screening procedures have been evaluated, including most of those that are the subject of present-day controversy. In selection of procedures for consideration it has been recognised that evaluation may be premature in some where knowledge or methods are known to be seriously deficient, and it may be of no immediate value in others already provided as a national service. Among the criteria thought to give a procedure priority for consideration are the following—

(a) the relative importance of the condition to be detected by screening;
(b) pressure, or the likelihood of pressure, on government departments to introduce a national service;
(c) indications that evidence related to a screening procedure has reached a critical stage;
(d) grounds for believing that there might be substantial benefit from a particular form of screening;
(e) a procedure is one which has not been considered recently in detail.

PROBLEMS IN SCREENING
Finally, it should be recognised that the preparation and implementation of screening policies present difficulties, some of which are quite formidable. Many of these problems arise in any large scale investigation of the kind needed to evaluate clinical procedures in general. However, some are unique to screening; for example, the ethical issues raised by an approach through ‘at risk’ groups. Moreover, those that are not unique are unusually troublesome because of the stringent evidence required in respect of benefits and risks. I shall discuss these problems in relation to the acquisition and interpretation of evidence, and to what, for want of a better term, will be referred to as managerial problems.

Acquisition and Interpretation of Evidence
Although investigations are needed in countless areas of clinical medicine where knowledge is incomplete, when the patient presents to the doctor in the usual way, medical practice must continue in spite of the deficiencies. But where the doctor initiates the examination, as in screening, research to validate the procedure is an essential preliminary to its general introduction in service. Since the research may require careful surveys of large populations over long periods it is very difficult to organise and control. Consider, for example, the difficulties involved in obtaining a complete picture of the natural history of bacteriuria, or of breast or cervical cancer.

But even where it is possible to by-pass some deficiencies of knowledge of natural history, as in the recent H.I.P. investigation of breast cancer screening in New York, the difficulties may be very great. For instance, it is unlikely on ethical and other grounds that it would now be acceptable to assign patients to ‘screened’ and ‘control’ groups, as was possible in the New York enquiry. And even in New
York, the strict identity of the groups could not be maintained, since control patients who asked to be screened were not refused. The difficulties are likely to arise increasingly in research on screening: patients cannot be assigned randomly to screened and control groups and where they are assigned, the distinction cannot be strictly maintained.

Another kind of difficulty is illustrated by current differences of opinion about the risks from exposure to radiation in mammography. It has been estimated that at a dose of two rads per breast the risks are significant (Mole, 1974), and that at eight rads per breast per year—the level of exposure which was general in the past and which is still used in some places—it is possible that the number of cases of radiation-induced cancer might equal the number of lives prolonged (not necessarily saved) by screening. However, these estimates are not accepted by some people with long experience of breast cancer and problems related to screening (British Breast Group, 1975). I mention this example, not only to illustrate the very different conclusions reached by experienced people on the same evidence, but also because of the difficulty of obtaining the evidence that would resolve the differences. In a matter of this kind all that can be attempted is to estimate the range within which the risk falls, but it is very difficult even to establish the range. The issue is particularly important in screening, where it is unethical to expose people who believe themselves to be well to an unspecified risk.

Table 1. Results of the New York investigation of breast cancer screening: five-year follow-up

| Age at death | No. of deaths from breast cancer | Reduction of mortality (per cent) |
|--------------|----------------------------------|----------------------------------|
|              | Controls | Study group |                                |
| 40-49        | 12       | 13           |                                |
| 50-59        | 34       | 16           | 53                              |
| 60-69        | 17       | 11           | 35                              |
|              |          |              | 63                              |

Remarkably perhaps, there are problems in interpreting evidence about screening as well as in acquiring it. Table 1 gives results from the New York investigation of breast cancer screening, and shows the number of deaths from breast cancer in screened (30,000) and control (30,000) groups of women of different ages after a five-year follow-up period. Three errors were sometimes made in interpretation of this evidence in this country.

(a) It was concluded that, because the difference between 'screened' and 'controls' was not statistically significant in women aged 60 to 69, there was no evidence of benefit in this age group. All that can be deduced is that the numbers
are too small to give a significant difference; so far as they go the results in those over 60 years of age are consistent with those for younger women.

(b) The results of screening in women aged 50 to 59 was interpreted by some as a reduction of breast cancer mortality among screened women. Without follow-up for a much longer period it shows only increased survival to five years. In a later H.I.P. report the improvement was maintained for a longer period; but I am concerned here only with the interpretation of the evidence.

(c) H.I.P. results were extrapolated to mean that, if extended to all women, a service of this type would reduce mortality from breast cancer by about a third. We are indebted to Knox (1975) for drawing attention to the fact that in a disease manifested over several decades, the effect of a screening programme over a few years on the death rate from breast cancer would be very small. In respect of the H.I.P. results over seven years he wrote: 'It must be remembered that the 30,000 women in the H.I.P. study group could be expected throughout their lifetime to generate about 2,000 breast cancers with about 1,200 deaths and that the seven-year saving of 38 cases (that is, the difference between the study group and the control group) will represent only about 3 per cent of all these deaths'.

Managerial Problems
An example of the class of problems referred to here is the difficulty that arises in the application of the concept of 'at risk' groups. It is obvious that the efficiency of screening would often be enhanced if it could be restricted to people shown to be at special risk; and the smaller the proportion of the population and the higher their risk, the greater would be the efficiency of the service. Moreover, there are areas of medical practice where the application of such knowledge presents no particular difficulty; for example, people can be told that they are doing something which increases the likelihood of disease, as in the case of smoking, when, as a result of the advice, they are able to reduce or eliminate the risk. It is also possible to inform patients that they may have a disease, when the diagnosis can be confirmed or excluded at once by investigation, as in the case of biopsy for suspected breast cancer. Neither of these conditions is usually met in screening practice. Consider, for example, the possibility of identifying women at high and low risk of developing breast cancer. Those on whom screening services were concentrated would need to be told that they were at increased risk, and would have to live for years, indeed for the rest of their lives, with the threat of a serious disease that in a large majority of cases would prove to be unfounded. There would also be some women with breast cancer among those at low risk, and it might not be easy to convince them that the preliminary assessment, which they would consider to be in error, made it desirable to deny them screening services from which they might have benefited.

Another and a very different type of managerial problem is illustrated by the question whether it was desirable to repeat in Britain the H.I.P. investigation of breast cancer screening conducted in New York. A case could be made for doing
so: there was, and indeed there is now, no other reliable evidence of the benefits to be expected from this type of screening, and it seems quite unsatisfactory to base future policy on so important a question on the results of a single trial in another country. Opinion was, and to some extent still is, divided on this matter, even among those who have given a great deal of thought to breast cancer screening and to the interpretation of the H.I.P. evidence. In the light of the fuller picture of the issues now available, it seems clear that a further investigation of screening over a few years, as in New York, would not resolve the problems of a national service which would probably have to be extended over a much longer period. But the decision was difficult when it was first made and—the point I wish to emphasise—it could not be reached solely on the basis of the scientific evidence that was available. It is in this sense that screening, in common no doubt with some other areas of medical practice, but more frequently and more acutely than most, presents managerial problems.

This article is based on a paper read at the College conference, 'Achievements and illusions—the value of some current practices', held in Coventry in September 1975

References
British Breast Group (1975) British Medical Journal, 3, 357.
Knox, E. G. (1975) Probes for Health, p. 15. Nuffield Provincial Hospitals Trust. London: Oxford University Press.
Mole, R. H. (1974) Proceedings of M.R.C. Conference on Breast Cancer.
Nuffield Provincial Hospitals Trust (1968) Screening in Medical Care. London: Oxford University Press.