MULTIPHASIC SCREENING: A REVIEW AND SOME PROPOSALS FOR RESEARCH

N. J. Glass

January 1975 WP-75-5

Working Papers are not intended for distribution outside of IIASA, and are solely for discussion and information purposes. The views expressed are those of the author, and do not necessarily reflect those of IIASA.

Introduction

This report is intended as an interim document, setting out as far as possible the current state of the argument on multiphasic health screening and suggesting a course of research that IIASA, given its in-house personnel and its possibilities for contacts with other organisations, might wish to follow. The literature survey in the first part of the study cannot lay claim to being complete (in so far as this is ever possible), but it seems unlikely that any major study has been omitted which is likely significantly to alter the conclusions herein. This document is consciously structured as a research prospectus and it is hoped to elicit comments to it on this basis.

Definitions

The U.S. Commission on Chronic Illness (1) defined screening as "the presumptive identification of unrecognized disease or defect by the application of tests, examinations or procedures which can be applied rapidly....A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and treatment." Further distinctions can be made between mass screening and selective screening of "high risk" or other groups and between one-shot screening tests (such as the PAP smear for cervical cancer) and multiphasic screening which normally includes "a medical history and physical examination and a range of measurements and investigations (e.g. chemical and haematological tests on blood and urine specimens, lung-function assessment, audiometry and measurement of visual acuity), all of which can be performed rapidly with the appropriate

staffing and equipment." (2) In automated multiphasic screening mechanical or electronic devices administer the tests and the results or data are introduced directly or manually into a computer which does all the necessary calculations and records and analyses the results (3).

Although the definition of screening quoted above stresses the distinction between screening and diagnosis, in practice it is not always possible to maintain such a rigid distinction. A physical health examination carried out by a physician, for example, will contain elements of both screening and diagnosis and Whitby has pointed out that "another effect of the development of high-capacity automatic laboratory equipment has been to make available to doctors, for screening purposes, the same investigations as are available to doctors for the investigation of patients. It is not always possible, therefore, to follow up an abnormal finding revealed as part of a screening programme other than by repeating the same measurement, this time as part of a diagnostic procedure." (2)

There has been over the past few years a good deal of thought given to the characteristics of an acceptable screening programme, i.e. an attempt to set up a check-list of those factors which are necessary (or perhaps only desirable) for a screening programme to be implemented. Although the utility of such a check-list, or rather the way in which it is open to misuse, is not difficult to demonstrate, the one proposed by Wilson and Jungner (4), for example, probably encapsulates best the thinking of epidemiologists on the evaluation of screening procedures. It provides a handy measure against which to compare the present state of multiphasic health screening.

The Wilson & Jungner principles are

1. The condition being sought should be an important health problem, for the individual and the community.
2. There should be an acceptable form of treatment for patients with recognisable disease.
3. The natural history of the condition, including its development from latent to declared disease, should be adequately understood.
4. There should be a recognisable latent or early symptomatic stage.
5. There should be a suitable screening test or examination for detecting the disease at the latent or early symptomatic stage and this test should be acceptable to the population.
6. The facilities required for diagnosis and treatment of patients revealed by the screening programme should be available.
7. There should be an agreed policy on whom to treat as patients.
8. Treatment at the presymptomatic, borderline stage of a disease should favourably influence its course and prognosis.
9. The cost of case-finding (which would include the cost of diagnosis and treatment) needs to be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuous process, not a "once-for-all" project.

History and Current Status

Thorner (5), in his critique of multiphasic screening, suggests that the concept of multiphasic screening for the detection of disease was born in the period immediately after World War II and after enjoying a brief flourish, appeared to have died out by the end of the fifties. It revived again in the mid-sixties, however, and had by the end of the decade become once more a live issue in discussions of medical care in the U.S. The notion of the periodic health examination has a longer history having had its advocates even in the nineteenth century. In 1922 the American Medical Association House of Delegates approved the idea of periodic medical examinations of "persons supposedly in health" (6).

Thorner attributes the first and second coming of multiphasic screening to technical developments that made the testing of large numbers of people feasible by simplifying the test procedure and reducing the cost per test. The existence of large-scale screening programmes for tuberculosis and syphilis after World War II provided a ready-made bandwagon upon which other tests could be placed e.g. for diabetes and led to the development of a number of demonstration projects throughout the U.S.

By the end of the fifties most of these multiphasic screening programmes had failed, a failure which Thorner attributes to the fact that the programmes were not properly integrated into the existing medical care system. The programmes, which were normally carried out by local health departments made little or no provision for diagnosis, follow-up and treatment and hence incurred the suspicion and resentment of private

doctors who alleged that they "dumped large numbers of disease suspects upon the private practitioners and provided no financing or facilities for diagnosis and treatment."

The periodic health examination which had traditionally been confined mainly to executive and managerial employees of corporations had shown no similar decline but in 1965,forty years after the A.M.A. statement, Grimaldi, having reviewed the various reports on such programmes, was forced to conclude that the question was still unsettled as to whether the examinations were practical when their yield was weighed against the time, cost, facilities, skill and energy required to provide them.

The development in the 1960's of multi-channel chemical auto-analysers and computer techniques as well as increased concern with chronic diseases led to a resurgence of interest in multiphasic health screening. Faced with problems of "physician shortage" relative to growing demands for health care, the prospect of using methods which would allow automated techniques and paramedical personnel to be substituted for expensive physician time was clearly an attractive one. A number of large programmes had continued in existence througout the period. Prominent among them was the Kaiser-Permanente programme, where multiphasic screening was embedded in a large prepaid health scheme. This scheme attracted particular attention due to the lead it provided in the use of automated techniques.

A survey in 1969 by the U.S. National Centre for Health Services Research and Development indicated that at that time there were about 150 Automated Multiphasic Health Testing (AMHT) programmes in operation in the U.S., the majority of them not receiving any

form of governmental financial support (7).

In Sweden AMHT was used by a group of six non-medical personnel to screen 89,000 persons in Varmland in the beginning of the sixties. This was to be followed up by further research programmes culminating in the trial screening of an entire county (about 250,000 inhabitants) in 1974 (8).

A number of small-scale trials have been carried out in the U.K., notably by Scott and Robertson in Edinburgh (9), by Holland and Trevelyan in London (10), and by Bennett and Fraser in Northumberland (11).

In Japan, the Toshiba Screening Programme provides an automated multiphasic health screening system for the 115,000 employees of Toshiba. In 1970 it was the only one of its kind in Japan. In Yugoslavia an ongoing collaborative project between the American NCHSRD and a number of Yugoslav health agencies is providing an experimental multiphasic screening programme in Montenegro.

Information on other examples of multiphasic screening is fragmentary. In Austria a feasibility trial of a national multiphasic screening programme was carried out in Vienna and Carinthia, in which 25,000 out of an invited 100,000 persons took part. The scheme is now being gradually introduced on a nationwide basis. A local scheme in Vorarlberg has also been reported, although its future relationship to the federal programme is uncertain.

It would therefore seem that multiphasic screening, and especially automated screening, is likely to become increasingly a candidate for health service resources.

The Case Against Multiphasic Screening

Much of the criticism of multiphasic screening (or the periodic health examination) has centred on the fact that while such testing discovers many abnormalities there is little evidence that such discovery leads to a better prognosis for the patient. After reviewing a series of studies reporting the experiences of patients who had undergone some form of early disease detection procedure Thorner (5) concludes "The evidence adduced by these studies for or against the effectiveness of multiphasic screening can hardly be considered definitive." Although many studies showed some improvement in morbidity and mortality of a tested group compared with a "control" population, none of the studies represented a properly designed randomised controlled trial and therefore considerable doubt must always exist as to whether the "control" population was really comparable.

Similarly Siegel in his review of the Periodic Health Examination (11), observes that there is no proof that populations receiving Periodic Health Examination (PHE) live longer, happier or healthier because of it, nor is there proof to the contrary. "PHE rests on the basic premise that discovering disease (or disease propensity) in the asymptomatic stage permits favorable intervention. Doubt is raised as to the validity of the premise as it applies to the prevalent, significant American adult diseases." Siegel suggests that if it is desired to persist with a policy of periodic health examinations, despite the lack of evidence of effectiveness, the policy should be modified to consist of the encouragement of Early Sickness Consultation for the majority of diseases for which pre-symptomatic detection is of no proven benefit. combined with periodic selective mass screening campaigns, using little or no

medical personnel, for the "relatively few amenable silent diseases." He does not explore in any detail, however, the resource consequences of the alternative programmes or the difficulties of encouraging "early sickness consultation."

Sackett (12) cites the early results from the Kaiser-Permanente trial of multiphasic screening as evidence for the ineffectiveness of such programmes. After several years of the programme these investigations were unable to determine any favourable health effect of the periodic health examination on women and only one group of men between the ages of 45 and 54 showed differences in disability and absenteeism. "Furthermore, these differences, while statistically significant are clinically unimpressive--only 3.9 % less disability and 1.3 % better attendance at work. The results of this study are quite sobering."

Schor et al. (13) in their examination of patients who had died and who had previously received a periodic health examination attempted to determine how often the examination had detected the subsequent cause of death. This, of course, is quite apart from the question of whether anything could have been done to prevent this. He found that, in all, the subsequent cause of death was only discovered in 51 % of the patients who died and the success rate was much higher, naturally enough, the nearer the PHE had been to the patients' death. This suggests that not only had PHE only detected about half of the causes of death but that, from the point of view of intervention, the utility of even these discoveries would be much reduced by the late stage at which they were discovered. Furthermore a study of matched living counterparts indicated that the same diseases that caused death were diagnosed with considerable frequency in those who did not die. The problem of dealing with

such "false positives" or "borderline" cases is another recurring problem in the evaluation of screening procedures.

A similar criticism is made by Sackett (12) who points out that most victims of coronary attack do not have clinically abnormal levels of serum cholestorol, blood pressure, triglycerides, uric acids or other risk factors; the number of victims with abnormal values for these coronary risk factors, despite their higher attack rates, are relatively few in number. Since, in his view, "the treatment of abnormal levels for the most prominent of these, blood pressure, does not appear to lower coronary risk, it must be acknowledged that the treatment of risk factors is not likely to have a profound impact upon the underlying burden of disability and untimely death."

The application of multiple biochemical screening on a routine basis came under fire from Ahlwin (14) and Barnett et al. (15). While reiterating the criticism that little evidence exists about the ability of physicians to influence the course of many of the abnormalities they discover through such screening, they also point out the ambiguity of many of these biochemical measurements from the point of view of clinical significance. Barnett cites a study in which calcium analyses were carried out routinely on approximately 12,000 patients. Since significance levels are normally set at the 5% level, approximately 600 were deemed to have abnormal results. The analysis is detailed in Table I.

Table I

Results and Follow-Up of Routine Calcium Analyses

Number of Patients	11,991	(100%)
Abnormal	600	(5%)
Significantly Abnormal	21	(0.21%)
% of abnormals		3.8%
Other not significant	539	(4.8%)
% of abnormals		96%
Diseases Found	23	(0.23%)
Diseases Treatable	14	(0.14%)

Source: Barnett et al.

Barnett points out that the discovery of these 14 diseases necessitated additional studies of 600 people, 571 of whom gave abnormal results because of laboratory errors, known diseases or for no reason ever found, and observes that "the amount of harm done to the 577 persons is not measured."

A point made by both Ahlwin and Barnett is that since biochemical tests are, in general, designed so that 5% of the results are termed abnormal, a screening programme involving a combination of tests administered simultaneously is likely to lead to a score of "abnormals" well in excess of 5%. In a situation where 12 constituents are measured one would expect that over half the patients would have at least one abnormal value and many of these will require follow-up and confirmatory tests. This leads into the question of the impact on the health care system of such screening.

Many critics have suggested that it is very unlikely either that over-stretched health services in the U.S. and Europe could provide sufficient manpower and facilities to carry out such testing or that the system could cope with the necessary follow-up,

diagnosis and treatment which the finding of such cases would imply. To the extent, of course, that multiphasic screening prevented a significant number of chronic diseases then in the longer run such screening might reduce the demand made by such diseases on the health services. In the short run, however, it would seem likely that such screening programmes would impose a net additional burden. Even in the longer run, it might well be that early detection lead to the patient requiring long-term maintainence therapy for an otherwise fatal disease, a result which, however, desirable in itself, is unlikely to lead to a reduced use of health services. On the basis of the preliminary results of the Kaiser-Permanente study, Thorner discerns an excess in the use of outpatient facilities by those patients receiving more screening tests over the "control"group of patients. The evidence, however, for this effect is limited.

One final problem with multiphasic screening may be noted and that is the reaction of physicians to the information provided. Bates and Yellin (16) found for only three out of 15 tests administered by a multiphasic screening programme did the patient's own physician carry out a confirmation more than half the time. When reasons for not doing so were examined, it was found that in over one-quarter of the cases this was because the results were either borderline or were unaccompanied by clinical manifestations. Bates & Yellin suggest that this is largely because of the high probability that the results would turn out to be a false positive in cases of diseases with low prevalence and that such unwillingness may therefore be quite "rational" but, of course, such behaviour necessarily reduces the utility of the screening programme.

Barnett cites a study in which screening profiles were carried out on an unsolicited basis for 400 patients. These

results were later presented to the attending physician after the patient was under treatment and the tests were scored on the basis of whether they were helpful, a hindrance (in the sense that they led to further fruitless studies), or neither a help nor a hindrance. Tests regarded as a hindrance occurred eight times as often as tests regarded as helpful.

The Case for Multiphasic Screening

Proponents of multiphasic screening (or periodic health examinations) rely on two sorts of arguments. The first attempts to show that such procedures are, in fact, effective, in the sense that they do lead to a reduction in disability, time-off work and mortality and that they do not involve an excessive burden on the health service. The second argument is that multiphasic screening is an essential element in a new system of medical care which ought gradually to replace the present system.

Grimaldi, after a review of previous studies of PHE which were largely inconclusive as to benefit, analyses data for three groups of General Electric workers. One group consisted of a random sample of middle management employees at a particular plant who had volunteered for a routine PHE which the company had been offering for many years. The second group was a random sample of non-participants and the third a similar group of employees from another plant where such examinations were not made available by the company. The groups were then compared on the basis of medical and surgical expense claims submitted to the company's insurance plan for a period of eight years.

The results were as follows:

1. The number of medical insurance claims per examined claimant increases with the time between examinations.

2. The smallest number of claims occurs during the year of examination.
3. The difference between the examined and unexamined, with respect to the average number of claims for claimant is negligibly small.
4. The medical expense per claimant increases as the time between examinations increases.
5. The average claim is greater for the unexamined and the difference tends to exceed significantly the cost per PHE and, in addition, occasions for payment for the treatment of coronary heart disease, circulatory disorders, malignancies and diabetes in the unexamined sample were somewhat higher than in the examined.

These results would appear to answer some of the questions raised by the critics of multiphasic screening and PHEs. In particular, the burden imposed upon the health services by this scheme would not appear to be great and there is indirect evidence of a consequent reduction in morbidity. It should be noted, however, that the study group volunteered to take a PHE while the controls either did not or could not. In this sense the groups may not be strictly comparable. It should also be noted that Grimaldi is principally interested in whether PHEs are a profitable activity from the firm's point of view. The lower health care expenses may simply represent false reassurance and do not tell us anything about eventual mortality. Finally, the caveat of Thorner should be borne in mind. After suggesting that the success of screening appears to be related to the medical care services already available to the population, he points out that in a study among persons in the lower socio-economic group with poor access to medical care and presumably

poor follow through after screening, the examinations seemed to make little difference.

Roberts et al. (17) compared the mortality experience of a large group of patients who had undergone one or more PHEs with that of groups of the U.S. population. In particular, given the socio-economic make-up of the study group, they compared the group's experience with that of professional, technical, administrative and managerial workers among the white, male population. Using this latter comparison, the mortality ratios of the study population were appreciably less than 1.0 in each category for which comparable U.S. data were available.

The Commission on Chronic Illness conducted a multiphasic screening in Baltimore in 1954. Subsequently a follow-up study was carried out twelve years later (18). The study showed that screenees and especially female screenees had a better survivorship than did individuals who had refueed screening and these differences persisted after adjustment for social class and for variations in history of chronic diseases or disability at entry to the study. However, in only one of the 14 age, race, sex combinations did the confidence limits for death rates fail to overlap, namely white women aged 40-49. Once again the problem of participation bias also affects the interpretation of the results.

The Kaiser-Permanente Medical Group, which was one of the pioneers of multiphasic screening, has been carrying out a trial of the procedure since 1964 (19) (20). The study group consists of a random sample of members of the Health Plan who have been encouraged to take advantage of the periodic multiphasic screening offered to all members, while the main control group consists of a random sample of members who have not been so

encouraged. The study group had, after the initial period, annual examination rates of 60% - 70%, while the rates for the control group were 20% - 24%. Seven years after the initiation of the urging effort, the average cumulated number of multiphasic health checks per subject was 3.5. in the study group and 1.3 in the control group. 17% of the study group had no examination during the period compared with 47% of the control group.

Analysis of the data through 1970 showed the following principal results:

1. There were no significant study-control group differences in utilisation of outpatient services, i.e. no overwhelming demand for additional outpatient services seems to have been generated by the increased study group exposure to Multiphasic Health Checks, nor has there been a notable reduction in utilisation.
2. In one of the four age-sex groups (men aged 45-54) a significantly higher rate of self-reported disability emerged after five years and persisted in the following bi-annual survey. In the other three groups (men and women aged 35-54) no significant difference had appeared after seven years.
3. There were no statistically significant difference in hospital utilisation between the study and control groups, although towards the end of the period utilisation rates appeared to be consistently higher among older control-group males and lower among females.
4. Causes of death among study and control population were divided, before the results were known, into two classes, a class likely to be reduced in persons taking periodic multiphasic checks and a remainder group. The results

are set out in table 2.

Table 2

Death and Death Rates Among Study and Control Groups 1965-71

	Death		Death Rates		Chi-Square
	Study	Control	Study	Control	Value
Potentially Post-ponable Causes	19	41	3.7	7.4	6.55*
Other Causes	164	176	31.9	31.8	0.00
All Causes	183	217	35.6	39.2	0.95

* $p < .025$

Source: Dales et al.

From this it can be seen that the class of deaths labelled "potentially postponable" was significantly smaller among the study group population, the greatest contribution to the difference being attributable to cancer of the colon and rectum and hypertension-hypertensive cardiovascular disease.

A study of screening by general practitioners in the U.K. by Bennett allowed him to estimate that the extra consulting time occasioned by the screened group (apart from the time required for the screening itself) was about 10%. Jungner and Jungner, in their Vaermland study (21) report on the disposal of screening subjects to the health care system. About 4% of those screened were referred to a doctor for diagnosis and about 0.3% were hospitalised. They remark that "this figure contrasts sharply with the fears of some advisors before the screening started." The significance of the figures is, however, difficult to interpret in the absence of measures of effectiveness.

The second line of advocacy of multiphasic screening sees it as the basis of a new method of organising health care which will move the focus of health care from treatment of symptomatic

disease towards prevention by the use of paramedical personnel and developments in medical technology. Garfield (22), for example, sees automated health testing as providing a means of regulating entry into the health care system which is fairer and more efficient than the pricing system. "Much of the trouble with the existing delivery system derives from the impact of an unstructured entry mix on scarce and valuable doctor time. Health testing can effectively separate this entry mix into its basic components:the healthy, the symptomless early sick and the sick. This clear separation is the key to the rational allocation of needed medical resources to each group.... The clear definition of a health care service, made possible by health testing, is a basic first step towards a positive program for keeping people well.... Whether or not one believes in the possibility of actually keeping people well, however, is now beside the point; this new health-care service is absolutely essential in order to meet the increasing demand for just this kind of service and to keep people from overloading sick-care resources."

Shapiro (23) puts it like this, "Initially A.M.H.T.'s goal was to aid in the detection of previously unknown disease. This has been expanded to include identification of patients with high risk for development of chronic disease and the initiation of health education to modify personal practices associated with adverse risk." The extent to which health testing can alter patients' behaviour in the long run is still unknown.

There would also appear to be certain ambiguities in the notion of health-testing as a regulator of entry into the health care service. Garfield argues that fees act as a deterrent to use of the service, but that some other regulator is necessary

and health testing could act as this regulator. The Kaiser-Permanente A.M.H.T. procedure lasts, however, somewhere between two and three hours, quite apart from any travelling and waiting time. Since individuals have time as well as money budgets, it would seem likely that this two to three hour time-expense would also act as a deterrent. That it does so is, perhaps, evidenced by the fact that despite the screening procedure being a money-free service to members of the Health Plan, only about 20% of members actually take advantage of it in a normal year. Even when a sample of members were subjected to intensive encouragement to take part, 17% did not do so once during a period of seven years. Whether such time-prices are a fairer way of regulating entry into the system than money-prices could presumably be decided if information existed on the type of people discouraged by either method. There seems no a priori reason, however, why time-price regulation should regulate entry more in line with "need" than money-price regulation.

Multiphasic Testing and the Health Care System

Questions about the place of multiphasic testing in the health care system have received a lot of attention and it has been frequently suggested that unless multiphasic testing is properly integrated with the rest of the health care system it will not succeed. Clark and Ariet (24), after reviewing successful AMHT set-ups, rank integration within the local health delivery system as their condition for success.

By "integration" is meant that the testing should operate

within a system of health care which provides for follow-up, diagnosis and treatment and that it should operate through, or with the cooperation of the patient's personal physician. Related to this is the provision that sufficient resources should exist to ensure that the results of the screening are acted upon where necessary. Otherwise there is the danger that, as Bates and Yellin discovered, many patients will not be seen by their physician following screening. "The complexity, discontinuity and fragmentation of medical care complicated by patients' misunderstanding, took their toll. This emphasizes the importance of close collaboration between screening unit and medical care system."

Some authors have suggested that the institution of AMHT is bound to fail unless health care delivery systems are reorganised so that doctors operate from large prepaid group practices which could afford, or justify, the capital investment and paramedical personnel required for Automated Multiphasic Health Testing and yet provide the continuity of care required (25) (26). It is also claimed that unless such a reorganisation takes place multiphasic health testing will not reach the medically underprivileged.

On the other hand the schemes currently in operation in Austria and Great Britain rely on general practitioners to carry out the screening procedures and, in general, imply no radical alteration in the structure of health care delivery.

Costs and Benefits

Very few studies exist which attempt to estimate the costs and benefits of early disease detection, a fact which is hardly surprising given the paucity of data on the outcome of screening

programmes. The study by Grimaldi mentioned already claimed that the excess of health insurance claims of an unscreened population over a screened one more than offset the cost of screening such a population.

The Kaiser Permanente study found only one of its four age-sex groups where disability, hospitalisation and mortality trends were all in the same direction (favouring the study group), namely the group of men aged 45-54 on first entry. The direct (medical) costs and indirect (income loss) costs associated with screening, outpatient services, hospitalisations, self-reported disability and mortality were computed and compared for the older men in the study and control groups. For this group the net "savings" favoured the study group older men every year during a seven-year period and was more than $800 per man entering the project in 1964, in favour of the study group, for the entire seven-year period. This sort of measure is, of course, open to the usual reservations about the use of income loss figures as measures of benefit.

A rather different impression emerges from a remarkable study by Forst (27) in which he analysed data from Periodic Health Examinations given to Navy and Army officers. Because of differences in the scope and frequency of PHE's between the two armed services he was able to make estimates of the cost and morbidity effects of changes in the scope and frequency of PHE's. He estimated that a shift from a strategy of giving a PHE worth $25 once every three years to that of giving one worth $100 annually could be expected to prevent about seven officers out of each 10,000 from joining the rolls of disabled retirees annually. Such a shift would cost $150,000 per head.

When this sum was compared with the cost of retirement benefits plus the cost of replacing officers retired with disability it was found that a replacement cost of $130,000 would be necessary to make these two costs exceed $150,000. The likely replacement cost was estimated to be less than one-fifth of $130,000. The data and assumptions in this study, as well as some of the regression relationships estimated, make one treat the results with caution but in sophistication and rigour it far excels any other study, although the sophistication and rigour may well be a product of the initial data problems.

A number of costing studies have been carried out by the Kaiser-Permanente workers (28) (29). They estimated that for 1967/68 the cost per automated multiphasic screening test at their Centre was $21.32, based on a monthly total of 2000 patients. They believed this cost to be four to five times less than the cost of providing an equivalent screening by non-automated methods. They also estimated that there was significant economies of scale in automated screening and suggested, for example, that if the number of patients were to rise to 3000 per month the cost might fall to $15 per screening.

Estimates were also made of the cost of detecting various abnormalities by multiphasic screening. These costs ranged from $408 for a suspected breast cancer to $1.55 for an apparent hearing defect. No estimates were made of the cost of confirming such presumptive cases.

There are some points to be made here. The costs estimated by Collen and his co-workers refer to the personnel and equipment for the multiphasic centre. They do not include any estimate

of patients' time. The Kaiser-Permanente multiphasic schedule takes up to three hours. If patients' time is valued on a marginal productivity basis using an average wage rate of $10,000 per annum, then such a valuation would effectively double the reported cost of multiphasic screening.

Secondly, as Collen points out, no analysis has been carried out "comparing the cost of multiphasic screening technics to alternative traditional methods for providing periodic health examinations." Austria has opted for a fairly traditional method relying on a primary physician and the costs do not appear to be regarded as excessive. Nor would there seem to be any analysis comparing the cost of adding extra tests under both systems.

Finally, very little has been said about the speed with which a multiphasic programme could be introduced given its large demands for personnel and facilities. Gelman (30) points out that to carry out annual periodic health examinations for the entire U.S. population, based on an estimate of 60 to 65 "family service" physicians per 100,000 population available for diagnostic services, each physician would have to perform roughly seven physical examinations a day. As she says "Who would then care for the sick?"

Conclusions

1. Multiphasic Screening or Periodic Health Examinations still appear to be seriously deficient when assessed against the checklist suggested by Wilson and Jungner. In particular it is not fully established that treatment at the presymptomatic borderline stages of the "diseases" discovered, favourably influence their course and prognosis. Moreover it is not certain that the

facilities required for diagnosis and treatment of patients revealed by the screening programme would be available and that case-finding in a given health care system would be a continuous process, and not a "once-and- for all" project.

A number of projects are going on, in addition to the Kaiser-Permanente one,to assess the effects of multiphasic screening (25), but until these projects have progressed further, it is difficult to say whether the cost of case-finding (defined in as comprehensive a manner as possible) would be economically balanced in relation to possible expenditure on medical care as a whole, since such a decision could only be taken with reference to the benefits of multiphasic screening relative to possible expenditure on medical care as a whole.

2. Multiphasic Health Testing can be viewed in two different lights:(a) as a multiple screening programme or (b) as the basis of an alternative method of delivering primary care.

Viewed in the first light, the important question becomes "Is multiphasic screening worth doing, in the sense that the outcome of multiphasic screening represents a more desirable use of limited medical resources than other programmes?

Viewed in the second light, the crucial question may well be "What would be the effect on present health care systems of an extension of the practice of multiphasic screening?" By this is meant not simply what resources would be required, but how would the pattern of health care utilisation alter as "prevention" became a more important characteristic of primary care, what changes would be likely to come about in the way in which primary health care was supplied and how would this affect other sectors of health within a limited budget. Such effects might be expected to vary depending upon the way the programme was organised.

3. There is very little information available, apparently, on a number of topics. The gaps which appear most evident to me are

(a) The lack of any systematic analysis of the different organisational structures through which multiphasic screening is currently being carried out. A formal structure to classify schemes by payment mechanism, physician participation, relationship to hospital, role of public health services, degree of automation, location and so on is missing and, especially for Europe, the data to flesh out the formal structure is also absent.

(b) Data on how much screening goes on in a given health care system is difficult to come by. Of course, any precise estimate would depend on distinctions made between screening and diagnosis. Nevertheless, judgements about the impact of multiphasic screening might well depend on the amount of screening going on outside such programmes (e.g. routine testing of hospital patients, expectant mothers, life-insurance candidates, specialist professions, etc.) and about trends in such activities.

4. There would appear to be some dispute about the possibilities of inserting a multiphasic screening programme into a health care system without making radical changes in the system.

5. The work on costing which has been carried out has been very limited and has generally not addressed itself to the total system, i.e. achieving participation, screening, patient time, follow-up, diagnosis and treatment. Nor has there been any attempt to assess the effects of adding or subtracting tests or altering the administration of the programme. In short, there has

been little costing of alternatives.

Research Proposals

I set out below a list of projects that IIASA might carry out. It is not intended, necessarily, that all of them should be done and the choice should take place in consultation with interested parties such as WHO, the health ministries of member states or other organisations connected with health care.

Survey

A sine qua non for any further work is a survey of what programmes, described as "multiphasic screening", are currently being carried on in a number of countries, both East and West. Our aim should be to characterise them by a number of relevant variables. Warshaw (3) sets out a number of "questions to be answered" in deciding whether to organise a multiphasic screening programme. Suitably amended, these questions would form the basis of a scheme of classification. An additional important piece of information would be the sorts of data which the programmes themselves collect and the extent to which this data is available in treated or untreated form.

The bases for classification would be

1. Target Populations and Diseases
2. Financing
3. Number and Location of Screening Clinics
4. Manpower
5. Selection of Tests
6. Volume of tests carried out and participation rates
7. Provision for follow-up, diagnosis and treatment
8. Management of the Scheme
9. Cost of the Scheme

Most of these classifications are fairly self-explanatory. Financing refers to the inclusion or exclusion of such screening programmes in relation to Social Security programmes, public health programmes, etc. and the extent, if any, to which patients must pay. Manpower refers to the relative use of medical and non-medical personnel. Management of the Scheme is intended to elicit information about the relationship between the screening programme and other health care services in addition to that provided by No. 7.

Some of this information may already be gathered for the U.S.A. by the National Center for Health Services Research and Development. For other countries, it is difficult to know even whether such programmes are in existence.

I would suggest such a survey, in effect as a form of feasibility study.

Health Service Impact

So far there has only been limited work on what effect the introduction of a multiphasic screening programme might have on the existing system of medical care, and how this effect might be related to the type of programme. In general, the worst fears of those who opposed such programmes, namely that the existing health care system would be swamped by large numbers of the crypto-ill does not appear to have been borne out. There has, however, been little analysis of why this is so. Were the critics simply wrong i.e. there are not large numbers of slightly unwell people or are there other explanations? Is it that the "acceptance" level was simply set at a level such as to ensure that there would be little impact on the health service, or were health service facilities already so strained that, in effect, a form of rationing ensured that the screening programme was not allowed

to make a major impact? The answers to these kind of questions would require close analysis of the experience and criteria of a number of screening programmes.

A related question is the way in which the institution of such programmes affects the use of primary care facilities, out-patient services and so on under various forms of health care organisation. Does the pattern of demand for physician's services and the use by the physician of, for example, laboratory facilities change? The data for such a study might be available from records of Social Security sickness claims,hospital records and so on.

Modelling Alternative Multiphasic Screening Programmes

Choosing a multiphasic screening programme implies some knowledge or prejudice concerning the effects of alternative strategies. As yet, however, there has been no attempt to construct a model which would predict both the yield of alternative multi-phasic screening programmes and the costs of such programmes.

Such a model would require, in the first instance, estimates of suspect and confirmed abnormalities for a range of individual tests and combinations of tests over time together with data on the proportions of confirmed populations likely to undergo various forms of treatment. This would enable one to make estimates of the consequences in terms of cases discovered and effects on the health care system of adding an extra test or set of tests to the programme.

It would also require estimates of the resources used up to supply a given set of tests and the marginal costs of adding extra tests to the range, as well the health care resources required for follow-up and treatment. Added variables here could be the organisational structure, the use of medical and non-medical personnel, of automated equipment, different types of clinic and

so on. It might also be possible to derive estimates of the interaction between organisational characteristics of the screening programme and the participation rate of the target population.

The data for such a study, or at least parts of it, should be in the possession of screening programmes. To move from this partial analysis of the effects of a screening programme to a comprehensive analysis or a cost-benefit analysis would require estimates of the effect of multiphasic screening on morbidity. The evidence for this, as has been said, is still inconclusive. But even with the partial model it should be possible to elicit estimates of the impact of a screening programme by making (probabilistic) assumptions of the effect of early discovery.

This proposal would overlap with the previous one but would not attempt to trace the effect on the health services in general of a multiphasic screening programme nor to concern itself with the impact of individual programmes.

Case Studies

The introduction or adoption of multiphasic screening or periodic health examination programmes takes place within a given social and institutional setting. The programme might be expected to affect some or all of the following: the target population, the population in general (in so far as the programme reduced expenditure on other health services or implied higher taxation or social security contributions), the medical profession, both in the community and in hospitals, public health administrations, health insurance institutions, manufacturers of relevant equipment.

Where such programmes have been introduced the choice of programme, its administration and so on are likely to have been as much

the consequence of the constraints represented by these various groups as of the benefits and costs of the screening programme itself. An analysis of the proposals, attitudes and reactions of these groups in relation to a specific screening programme might be expected to throw light on some of the constraints operating on the introduction of such screening programmes. Many of these constraints would, of course, be confined to particular settings but the attitudes of groups should also be found to have certain general characteristics.

A useful starting point might be an analysis of the introduction of the Austrian multiphasic screening programme.

REFERENCES

[1]Commission on Chronic Illness. Chronic Illness in the U.S., vol. 1, Harvard University Press 1957.

[2]Whitby, L. G.; "Screening for Disease - Definitions and Criteria," The Lancet October 5, 1974.

[3]Warshaw, L.; "Principles of Screening and Types of Programme," Medicine Today, Vol. 5, No. 2; 1971.

[4]Wilson, J.M. and Jungner, G.; Principles and Practice of Screening for Disease; WHO 1968.

[5]Thorner, R.; "Whither Multiphasic Screening," New. Eng. Jour. Med. Vol. 280, No. 19.

[6]Grimaldi, J.; "The Worth of Occupation Health Programs," J. Occup. Med.; No. 7; August 1965.

[7]Thorner, R.; "The Status of Activity in Automated Health Testing in the U.S," in AMHT. C Berkley (ed) Engineering Foundation, New York, 1971.

[8]Malmstroem, G.; "AMHT in Europe," in Automated Multiphasic Health Health Testing. C. Berkley (ed) Engineering Foundation, New York, 1971.

[9]Scott, R. and Robertson, P.; "Multiple Screening in General Practice," British Med. Jour. 1968, 2, 643.

[10]Bennett, A. and Fraser, I.; "Impact of a Screening Programme in General Practice: A Randomised Controlled Trial," Int. Jour. Epid. 1972, Vol. 1; No. 1, p; 55.

[11]Siegel, G.; "An American Dilemma - The Periodic Health Examination," Arch. Envir. Health; Vol. 13, Sept, 1966.

[12]Sackett, D.; "Periodic Health Examination and Multiphasic Screening," CMA Journal; Vol. 109, December 1, 1973.

[13]Schor, S. et al.; "An Evaluation of the Periodic Health Examination," Annals of Internal Medicine, Vol. 61, No. 6, December, 1964.

[14]Ahlvin, R.; "Biochemical Screening - A Critique," New Eng. Jour. Med. Nov. 12, 1970.

[15]Barnett, R. Owen, H. and Schoen, I.; "Multiphasic Screening by Laboratory Tests - An Overview of the Problem," AJCP. Vol. 54, September, 1970 (Suppl.)

[16]Bates, B and Yellin, J.

[17]Roberts, N.J. et al.; "Mortality among Males in Periodic Health Examination Programs," New. Eng. Jour. Med. July 3 Vol. 281, No. 1, 1969.

[18]Kuller, L. and Tonascia, S.; "Commission on Chronic Illness Follow-up Study," Archives of Environmental Health, Vol. 21, Nov. 1970.

[19]Dales, L. Friedman, G. and Collen, M. F.; "Evaluation of a Periodic Multiphasic Health Checkup," Meth. Inform. Med. Vol. 13, No. 3,1974.

[20]Cutler, J. et al.; "Multiphasic Checkup Evaluation Study 1. Methods and Population," Preventive Medicine 2, 197 - 206, 1973.

[21]Jungner, G. and Jungner, I.; "Chemical Health Screening" in Sharp, C. and Keen, H. (eds) Presymptomatic Detection and Early Diagnosis, Pitman Medical, London, 1968.

[22]Garfield, S.; "The Delivery of Medical Care," Scientific American Vol. 222, No. 4, April, 1970.

[23]Shapiro, S.; "The Health Care Structure and Automated Multiphasic Health Testing Services," in Provisional Guidelines for A.M.H.T.S. Vol. 3, N.C.H.S.R.D. Rockville, Maryland U.S., 1970.

[24]Clark, E. M. and Ariet, M.; "Automated Health Testing and Medical Care," Jour. Florida Med. Assocn. Vol. 59/10, 1972.

[25]Shapiro, S.; "AMHT - Efficiency of the Concept," Hospitals, Vol. 45, March 1, 1971.

[26]Badgley, R. and Wolfe. S.; "Public Policy and AMHTS," in Provisional Guidelines for A.M.H.T.S. Vol. 3, N.C.H.S.R.D. Rockville, Maryland USA, 1970.

[27]Forst, B.; "An Analysis of Alternative Periodic Health Examination Strategies," in Niskanen, W. and Wisecarver, D (eds) Benefit-Cost and Policy Analysis 1973, Aldine Publishing Co., Chicago 1973.

[28]Collen, M. et al.; "Cost Analysis of a Multiphasic Screening Programme," New. Eng. Jour. Med., Vol 280, No. 19, 1961 pp. 1043-45

[29]Collen, M. et al.; "Dollar Cost per Positive Test for Automated Health Testing," New. Eng. Jour. Med; Vol. 283, No. 9, Aug. 27, 1970.

[30]Gelman, A.; "Automated Multiphasic Health Testing," Public Health Reports, Vol. 85. No. 4, April, 1970.