SCREENING METHODS IN HEART DISEASE

PROBLEMS AND PITFALLS IN AVAILABLE METHODS

By Geoffrey Rose

There are signs in many countries of a wave of rising enthusiasm for mass cardiovascular screening. As doctors we were all trained to be happy when we diagnose a case of previously unsuspected disease: diagnosis is "a good thing", and we work hard to make as many diagnoses as possible.

In recent years the epidemiologists, like travellers, have set out through the hospital gates to explore the medically unknown territory of the general population. Now they are returning, bringing amazing stories about vast numbers of undiagnosed conditions: for every one case that doctors know about, they say that there are several more—hitherto unsuspected—cases of hypertension, and diabetes, and hyperlipidaemia, and coronary heart disease, and cardiomyopathy—waiting out there, undiagnosed.

Many clinicians have not yet comprehended the full implications of what these returned surveyors are reporting. But growing numbers of others can be seen setting out from hospital, their bags hastily packed with the familiar diagnostic tools, to make the most of these vast undeveloped diagnostic resources. Before we encourage their efforts, or decide ourselves to join them, we should perhaps be wise to ask a few questions concerning the real purpose of the expedition, and the suitability for this new venture of our traditional diagnostic tools; we should be aware of the problems and pitfalls, and the purpose of this short paper is to highlight the main ones.

BIOLOGICAL VARIATION

This affects most screening measurements. Even under standardised conditions, for example, the within-subject standard deviation of diastolic blood pressure averages about 7 mm. (Armitage et al., 1966). (For systolic pressure the variability is proportionately rather less.) If we take a single measurement to characterise an individual, then we are taking a random sample of 1 from a wide distribution: for all we know, it could easily be as much as 10-15 mm. above or below that person's true mean. Some subjects—whom we cannot identify in advance—have an even greater liability, with standard deviations as much as 20 mm.; in them, the uncertainty is correspondingly wider. This means that we are not usually justified in raising an alarm after a single screening examination: high values must be checked by repeated measurements, made on different occasions.

OBSERVER VARIATION

If two observers have a systematic difference of as little as 5 mm., then one is likely to find twice as many cases of "hypertension" as the other. If two laboratories have a systematic difference of 15 mg./100 ml. in estimating blood cholesterol levels, then one may find twice as many cases of hypercholesterolaemia. In practice, systematic differences are often much larger than in these two examples.

The exercise electrocardiogram is widely used to screen for evidence of early myocardial ischaemia. Blackburn et al. (1968) recently submitted 38 records from asymptomatic subjects for independent judgement by 14 leading cardiologists. The most pessimistic among them classed 58% as "abnormal"; at the other extreme, another observer called only 5% abnormal".

Epidemiology Department, St. Mary's Hospital Medical School, London.

This situation is tolerated—perhaps regrettably—in hospital practice only because we tend to function there as individual clinicians, each making unconscious adjustments of his management criteria in relation to his own personal diagnostic standards. In screening, this is not usually possible.

One response to the problem of observer variation in screening is to develop automated devices. These carry their own dangers, one being that an impressive collection of flashing lights and eleotronic circuitry may so dazzle us that we omit to ask some basic questions on validity and repeatability of the results.

The ElectroCardioAnalyzer* is a portable electrocardiographic screening device, designed to provide an immediate separation of subjects into "normal" and "suspect abnormal", according to pre-set measurement criteria for the various components of the complexes. My colleague, Dr. Christie, and I have recently evaluated the instrument in a screening study in factories; each subject also underwent conventional electro-cardiography, defined Minnesota Code criteria being used as the reference standard. Table 1 shows some of the results. The ElectroCardioAnalyzer appeared to have 12% false-negative results—a rather alarming figure. In screening, however, it is a mistake to lump all "true positives" together as though they were a single class: some are much less positive than others. In this instance a case-by-case review indicated that among the 2595 subjects there were only 7 "serious" false-negative —a figure that is probably acceptable. On the other hand, there was no escape from the estimate of around one quarter of subjects wrongly classified as positive.

This example is given as an illustration of the need for an objective numerical evaluation of every major screening method. Unfortunately this means a lot of work.

DIAGNOSTIC AND PROGNOSTIC SIGNIFICANCE IS DIFFERENT OUTSIDE HOSPITAL

Clinical cardiologists usually only see patients in whom heart disease is already strongly suspected. If an electrocardiogram in a man with a history of central chest pain reveals a rather prominent Q-wave, this strongly supports a diagnosis of myocardial infarction: but when recorded at a routine medical examination, a Q-wave of just the same dimensions has a very different diagnostic significance. Similarly, in hospital practice left bundle branch block is generally regarded as ischaemic, but in prospective epidemiological studies it has unexpectedly proved difficult to demonstrate any insignificant association with excess risk of coronary heart disease.

The same phenomenon applies to a wide range of physical findings: wherever a particular finding may represent either the result of disease or the extreme of physiological variation, then its significance will be much greater in hospital (where the ratio of abnormal to normal is high) than in the population (where the ratio is lower). When we enter the world of screening, we need to be re-educated in the interpretation of familiar observations.

INTERFERING IN THE LIVES OF UNCOMPLAINING INDIVIDUALS

In hospital practice the patient comes to us with a complaint, which establishes as **a priori** ground for

* Humetries Corporation of America.

giving treatment. Generally the case is further strengthened by a high level of confidence in our diagnosis, and by the serious prognosis of most cardiovascular diseases seen in hospital. But in all these respects the screening situation is different: it is we who seek out the patient, not vice-versa; because the tests have to be simpler, the level of diagnostic confidence is lower, and the prognosis is not nearly so unfavourable since we are dealing mostly with earlier or milder forms of disease.

These problems are illustrated by the situation in regard to screening for angina. A standardised questionnaire was developed (Rose, 1962) for research use in prevalence surveys—a very different objective from the characterisation of individuals in screening. The questionnaire is nevertheless attractive because it is cheap and simple to use: it can be administered by non-medical staff in an average of less than 1 minute per subject (or, in many populations, it can be self-administered). In industrialised countries the prevalence of "angina-positives", as defined by the questionnaire, is around 4% in middle-aged men, and most of them have not been previously diagnosed. Their risk of subsequent major disease or death is similar to that of men found to have S-T/T abnormalities in the electrocardiogram.

Thus in the questionnaire and E.C.G. we have two cheap tools available for immediate application in mass screening, capable of identifying sizeable sub-groups of the population with a much-increased statistical risk of myocardial infarction and death. Before advocating a screening policy we must, however, note two important points:

1. In absolute terms the prognosis for cases detected at screening is very much better than for hospital patients with angina or myocardial ischaemia. We have found (Rose, 1971) that 7 years later about 80% of them are still alive and well.
2. We have no evidence that their admittedly increased risk can be reduced by earlier intervention. Maybe it can: but at present the evidence is not there.

Similar considerations apply to most conditions diagnosed by screening. Diagnosis is not always a good thing. In screening it costs a lot of work and causes a lot of worry. Before advocating a screening policy we must require an objective evaluation of our diagnostic tools; and we must be aware that in both prognosis and therapeutic response, the cases detected at screening may behave very differently from their hospital counterparts.

REFERENCES

1. Armitage, P., Fox, W., Rose, G., Tinker, C. M.: "The variability of measurements of casual blood pressure II Survey experience." Clin. Sci., *30*, 337-344, 1966.
2. Blackburn, H.: *et al*, "The exercise electrocardiogram: differences in interpretation." Report of a Technical Group on exercise electrocardiography. Amer. J. Cardiol., *21*, 871-880, 1968.
3. Rose, G.: "The diagnosis of ischaemic heart pain and intermittent claudication in field surveys." Bull. Wld. Hlth. Org., *27*, 645-658, 1962.
4. Rose, G.: "Predicting coronary heart disease from minor symptoms and electrocardiographic findings." Brit. J. Prev. Soc. Med., *25*, 94-96, 1971.