and other dangerous substances may be taken unwittingly even in cold and cough remedies. A mandatory warning symbol such as a steering wheel, as suggested by Havard,[3] would be helpful on both prescription and over-the-counter preparations.

Drug abuse

Evidence about the accident risk associated with drug abuse is sparse. Hard-drug takers are underrepresented in accidents, presumably because they tend not to want to drive and probably cannot afford it. But effects such as euphoria and changed perception are incompatible with safe driving, and for cannabis there is some evidence of risk both from experiments and from accident surveys.

The study by Glauz and Blackburn[5] suggested a three-and-a-half times increased risk. A small and unrepresentative survey using more sensitive methods found evidence of cannabis in nine out of 66 dead drivers and motorcyclists, all under 30; but there were various uncertainties.[10] Evidence from surveys, however, is complicated by the fact that cannabis users tend to take other drugs, including alcohol, and to be untypical in various ways. Moskowitz made a comprehensive review of the experimental evidence.[11] He concluded that even at fairly low doses cannabis affects the perceptual elements in driving, probably through the central information-processing system, impairing concentration and judgment and the capacity to respond to signals and potential dangers.

The multiple drug abuse that is becoming more widespread might be particularly hazardous. Moreover, addicts being treated with methadone commonly drive and are known to have accidents, but the degree of risk has not been investigated. Though drug abuse cannot be a large cause of accidents addicts should clearly be warned of the dangers of driving and, whenever possible, individually discouraged from doing so.

I am grateful to the following for helpful discussion and comment: Professor R E Allsop and Mr R L Moore, transport studies group, University College London; Dr A B Clayton and Dr G M Mackay, accident research unit, University of Birmingham; Dr Griffith Edwards, addiction research unit, Institute of Psychiatry, London; Dr J D J Havard, British Medical Association; and Dr P A B Raffle, London Transport Executive.

References

1 Betts, T A, Clayton, A B, and Mackay, G M, *British Medical Journal*, 1972, **4**, 580.
2 Clayton, A B, Harvey, P G, and Betts, T A, *Psychopharmacology*, 1977, **55**, 9.
3 Havard, J D J, in *Medical Aspects of Fitness to Drive*, ed A Raffle, 3rd edn, p 43. London, Medical Commission on Accident Prevention, 1976.
4 Smart, R G, *Journal of Accident Analysis and Prevention*, 1976, **8**, 33.
5 Glauz, W D, and Blackburn, R R, paper presented to third international conference on drug abuse of the International Council on Alcohol and Addictions, London, 1975.
6 Storie, V J, paper presented to fifth international conference of the International Association for Accident and Traffic Medicine, London, 1975.
7 Sabey, B E, paper presented to joint conference of the American Association for Automotive Medicine and the International Association for Accident and Traffic Medicine. Michigan, Ann Arbor, 1978.
8 Kilbrick, E, and Smart, R G, quoted by Harvard.[3]
9 *British National Formulary 1976-78*, p 23.
10 Teale, J D, *et al*, *Journal of Forensic Science Society*, 1977, **17**, 177.
11 Moskowitz, H, *Accident Analysis and Prevention*, 1976, **8**, 21.

Epidemiology for the Uninitiated

Screening

GEOFFREY ROSE, D J P BARKER

British Medical Journal, 1978, **2**, 1417-1418

Screening patients for preclinical disease is an established part of day-to-day medical practice. Routine recording of blood pressure, urine testing, and preoperative chest *x*-ray films may all be regarded as screening activities; and the introduction of automated laboratory analyses has increased the scope of screening offered to patients in hospital and domiciliary practice.

Extending this activity to large-scale services for people who have not requested medical aid has been the subject of much recent discussion and controversy. It places the doctor in a new role, whereby it is he who seeks out patients and recommends treatment rather than the patients consulting him. Doctors therefore have a special obligation to ensure that screening is beneficial. To this end, there are three questions which must be answered and for which epidemiological data are required.

Department of Medical Statistics and Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT
GEOFFREY ROSE, DM, FRCP, professor of epidemiology

Faculty of Medicine, University of Southampton, Southampton General Hospital, Southampton S09 4XY
D J P BARKER, PHD, MRCP, reader in clinical epidemiology

Does earlier treatment improve the prognosis?

People with asymptomatic diabetes have a reduced life expectation. Nevertheless, two randomised controlled trials have failed to show that the prognosis is improved by treatment. In the Bedford study mortality after 10 years in a group of patients treated with tolbutamide was similar to that in a group treated with a placebo. In an American trial the cardiovascular mortality rate after eight years' follow-up was said actually to be higher in patients treated with tolbutamide or phenformin, although some doubts have recently been raised about these findings. There is, therefore, no indication that large-scale screening for asymptomatic diabetes is merited. This example illustrates that the outcome of screening must be judged in terms of its effect on mortality or illness, and not in terms of restoration of biochemical or other test results to normal.

In an American trial of breast cancer screening women were randomly allocated to a study group, which was offered annual screening, and a control group, which was not. Among the results was that the five-year survival rate of breast cancer cases detected by screening was 83% compared with 58% among cases in the control group. Such a comparison can be made only if there is allowance for the fact that earlier diagnosis will, of itself, increase the interval between diagnosis and death and thus improve survival rates over a short period. The figures quoted are so adjusted, and allow for the so-called "lead time"—the interval between the early diagnosis achieved by screening and the time when the disease would have been diagnosed without screening.

A further difficulty in this comparison of survival is that annual screening tends to identify cases of long duration, with a benign course, rather than cases where evolution is swift and fatal within a short period. This bias, however, does not apply to the main result of the study, which showed a reduction in the total number of deaths from breast cancer among the group offered screening as compared to the control group.

One facet of the effectiveness of treatment of asymptomatic disease is the definition of cases to be treated. In a previous article we showed that, since variables such as blood pressure have a continuous distribution within populations, there are no ready criteria for distinguishing "a case" from a normal person. Mild forms of the disease greatly outnumber severe forms. Somewhere towards the upper end of the blood pressure distribution there is a level above which treatment of asymptomatic cases will lead to a reduction in the frequency of strokes, heart failure, and renal damage. This level has yet to be defined precisely, and may vary from one population to another, but, for the purpose of screening, a case must be defined in terms of an operational level in so far as it can be determined from all the evidence—particularly that from clinical trials.

How valid and repeatable is the screening test?

Because a screening test must be inexpensive and easy to perform, it is not usually the best diagnostic method for a disease. In screening, therefore, it has to be accepted that some cases will remain undetected. The *validity* of a screening test is measured by comparing its performance with a reference test, and table I shows the outcome of such a comparison. Validity is compounded of *sensitivity*, derived from a/(a+c), and *specificity*, d/(b+d), which were discussed in our article published on 14 October, p 1070.

TABLE I—*A contingency table relating the results of screening and reference tests*

Screening test	Reference test	
	Positive	Negative
Positive	a	b
Negative	c	d

For variables such as blood pressure, whose distribution is continuous and unimodal, a rise of the threshold separating "normal" and "abnormal" people will increase specificity at the price of a reduction in sensitivity. The competing needs for high sensitivity and specificity must be balanced. A high sensitivity takes priority if a false-negative error is serious, as in screening for choriocarcinoma in women with a history of hydatidiform mole. High specificity is necessary when false-positive errors must be avoided—either because of the needless discomfort and anxiety to individuals, or because of resources required for further investigation.

In addition to its sensitivity and specificity, the performance of a test is measured by the *predictive value* of a positive or negative result. For a positive result this is given by a/(a+b), which represents the likelihood of a person with a positive test having the disease. When a disease has a low prevalence the proportion of true negatives (b+d) in the population in relation to true positives (a+c) is greater than when prevalence is high. Hence the proportion of false-positives (b) will be greater in relation to (a). The predictive value of a positive result must therefore fall as prevalence declines. Understanding this point is of practical importance, for new diagnostic tests are usually first tested in hospitals or clinics, where prevalence is high. Despite satisfactory levels of sensitivity and specificity these tests may be disappointing when applied to the general population, because the yield of false-positives is too great. Table II shows results from a breast cancer screening programme, using palpation and mammography, where the sensitivity was 67% and the specificity 98%, yet the predictive value of a positive screening test was only 20%.

TABLE II—*Results of a breast cancer screening programme using palpation and mammography*

Screening	Breast cancer		Total
	Present	Absent	
Positive	127	497	624
Negative	63	19 313	19 376
Total	190	19 810	20 000

Sensitivity = 67% (127/190); specificity = 98% (19 313/19 810); predictive value = 20% (127/624).

Assessing a screening test requires not only a comparison with a reference test but also measurement of the test's *repeatability*, which shows the extent to which a single screening measurement may be taken as a sufficient guide to action. Where subject variation, observer variation, or measurement errors are large, repeatability will be reduced and a single test result may be unacceptable.

What are the yields of the screening service?

The yield of a screening service is measured by the number of cases identified whose prognosis is improved as a result of their early detection. This must be related to the total number of tests performed. Theoretically, the yields of screening may be improved by restricting it to high-risk groups, as has been suggested in the screening of infants for developmental and other abnormalities. But identifying relatively small high-risk groups among whom most cases will be found is rarely feasible.

Although antenatal screening of women is widely accepted and practised, other forms of screening have proved less acceptable. The use of cervical cytology has been high in upper social groups, but low in social classes 4 and 5—where the disease has a higher incidence. Much has still to be learnt about methods of improving the acceptability of population screening, both in terms of changing people's attitudes and of the screening techniques that may be employed. The use of postal questionnaires, for example, requires further exploration. They are a cheap and acceptable technique in screening for hypothyroidism in high-risk groups, and there are postal "do-it-yourself" techniques which enable women to obtain their own specimens for cervical cytology.

Ultimately, the yields of a screening service have to be balanced against the costs, in terms of staff and facilities, for screening and making the confirmatory diagnoses. For breast cancer screening it has been found that identifying one case requires examining 170 women by palpation and mammography and taking nine biopsy specimens.

Eventually this series will be collected into a book and hence no reprints will be available from the authors.