Screening Tests in Mass Surveys and Their Use in Heart Disease Case Finding

By David D. Rutstein, M.D.

With illustrations by Ernest Craig, M.D.

Wherein certain principles governing screening tests for case finding in the general population are defined and the heart disease case finding studies in Boston are evaluated.

The combined experience from mass surveys for such diseases as tuberculosis, diabetes, syphilis, cancer and heart disease has made it possible to define in a preliminary way certain principles which govern the use of screening tests in case finding in the general population.

These principles are general and apply to the use of screening tests in mass surveys for any disease. These will be discussed with particular reference to the practical use of the 70 mm. photofluorographic film for heart disease case finding by health departments, heart associations and other interested agencies, based on the data collected in the Boston heart disease case finding studies.1-2 No attempt will be made to outline the administrative details of the organization and establishment of a mass survey, since these have been well summarized in recent publications.3-10

It might be assumed that a health inventory, consisting of a complete history and physical examination, performed generally throughout the population at annual or other suitable intervals would discover disease at an early stage and assure maximum benefits of prevention and therapy. Ideally this may be true, but actually it is impractical for many reasons, including the limited number of practicing physicians, lack of orientation of undergraduate and postgraduate medical education toward early diagnosis and the preservation of health, the traditional attitude of the average layman who seeks medical attention only during illness, and the enormous expense which must be borne by the individual or, failing that, by society.

There is therefore a need for simple and inexpensive screening tests to be used in mass surveys.* Screening tests are not diagnostic but are procedures which sort out those who probably have abnormalities from those who probably do not. Mass screening consists of the application of screening tests rapidly and economically to large population groups, to sort out those who probably have abnormal conditions and refer them for diagnosis and, if indicated, for further medical care.

Case finding through mass surveys among well people should not be confused with the extensive case finding performed by the average physician in individuals who come to him after the onset of symptoms and in discovery of disease other than that suggested by patients' symptoms. The mass survey supplements his usual case finding activities by referring to him early asymptomatic disease at a stage when therapy would probably be most effective.

Evaluation of a Screening Test

Before considering the use of a test for screening purposes in a mass survey, evidence must be available that:

1) The test is reliable. (a) The technical

*The definition which follows is that of the Committee on Early Detection and Screening, Section on Evaluation of Scientific Data, National Conference on Chronic Disease: Preventive Aspects.
features must be such that reproducible results can be obtained by technicians of average training. (b) Specimens collected at different times under the same conditions from the same individual must yield similar results within a narrow range of experimental error. Appropriate methods for determining such reliability are readily available.10-12 (c) Properly qualified personnel must be available for interpretation of the test.

In the Boston Heart Disease Studies, the 70 mm. photofluorographic film is a well standardized procedure7 which can be performed by average technicians. One unsolved problem still remaining is the exposure of the film at a specific phase of the respiratory cycle to avoid fluctuations in heart size due to changes in intrathoracic pressure during respiration.

The marked differences in yield of abnormal films in the two Boston studies1,2 were at least in part due to differences in training and experience of the readers in the interpretation of x-ray shadows of the heart and great vessels. It is evident that the success of a heart disease survey in a particular community will depend on the availability of properly qualified experts.

(2) The test is valid. The validity of the test is measured by the rate at which the result of the test is confirmed by an acceptable diagnostic procedure. Four possibilities exist when the results of the screening test are compared with those of the subsequent standardized diagnostic test: (a) "true positives": those who are selected by the screen and are also found to be positive by the diagnostic procedure; (b) "false positives": those selected by the screen who are found to be negative by the diagnostic procedure; (c) "true negatives": those who are not selected by the screen and are found to be negative by the diagnostic procedure; (d) "false negatives": those not selected by the screen but are found to be positive by the diagnostic procedure.

In order to be valid, a screening test must separate the "true positives" from the "true negatives," and keep to a minimum the number of "false positives" and "false negatives."

A mass survey divides a population into two groups, one in which the disease is likely to be present, the other in which the disease is likely to be absent. The first of these groups contains the "true positives" and the "false positives," the second the "true negatives" and the "false negatives." If the screen selects too many "false positives," an unnecessary burden is imposed on physicians and diagnostic facilities and the individual is put to needless anxiety and expense. On the other hand, if too many "false negatives" escape the screen, serious harm is done because these individuals are falsely reassured, since they are not selected by the screen and not referred for diagnostic examination. Figure 1 is an attempt to illustrate these theoretic concepts.

Special studies must be conducted to determine the validity of a screening test, since validity cannot be determined from the data usually collected in a mass screening survey. Thus, for example, in order to determine the validity of the 70 mm. film as a screening test for heart disease, in addition to the calculation of the ratio of the true positives to all positives, it is necessary to perform a complete diagnostic examination on a representative sample of those with a negative screening test to determine the ratio of the false negatives to all negatives. Such data would make it possible to determine how closely the results of the screening test approximate the true incidence or prevalence of heart disease in the study sample.

In the Boston heart disease case finding studies, the majority of those whose films were read as abnormal and who returned for examination were found to have clinically significant heart disease. Thus the number of false positives does not invalidate the practicability of the 70 mm. photofluorographic film as a heart disease case finding procedure. On the other hand, the percentage of false negatives is unknown because funds for the Boston studies did not provide for follow-up and complete examination of an unselected sample of those whose films were read as negative. Thus the ratio of false negatives to all negatives could not be calculated in the Boston studies and it therefore cannot be determined how well the screening test approximated the actual prevalence of heart disease in the Boston studies. However these studies were not designed to measure the validity of the test as defined, but rather to
Fig. 1. The color of the figures indicates the presence (red) or absence (green) of the disease. The screen has round holes and separates the population into two groups, square figures and round figures. The square figures are the "true positives" and "false positives." The "true positives" are the square red figures, those who are selected by the screen and who have the disease. The "false positives," the square green figures, are those who are selected by the screen but do not have the disease. The round figures pass through the screen. They are the "true negatives" and the "false negatives." The "true negatives" are the round green figures who are not selected by the screen and do not have the disease. The "false negatives" are the round red figures who escape the screen but do have the disease. (The use of color in this illustration is made possible by a grant from Wyeth, Incorporated to the publication fund of the American Heart Association.)

answer the practical question indicated in the first paragraph of the first paper of this series: "Is it possible, by means of the 70 mm. photofluorographic film as taken for tuberculosis, to identify in the general population a significant number of individuals with previously unknown heart disease who will be benefited by such identification?"
If a screening test is found to be theoretically valid, there still remain the practical problems of follow-up for return for diagnosis of those selected by the screen, provision of complete diagnostic facilities and referral to medical supervision of those who are found to have the disease. Figure 2 is an attempt to illustrate these concepts.

The Boston heart disease case finding studies demonstrated the practicability of performing a heart disease survey with the 70 mm. photofluorographic tuberculosis survey film under circumstances of cooperation among the many community agencies concerned with the organization and conduct of the studies (private physicians, medical societies, visiting nurse association and the agencies responsible for the tuberculosis survey), and in which complete diagnostic facilities were established. It is not known whether this survey method is practicable if diagnostic facilities are not established, since the 70 mm. photofluorographic film in many cases may provide only suggestive evidence of cardiac abnormality which will be of little assistance to the general physician in making a definitive diagnosis of heart disease, and may make it difficult for him to justify referral of individuals with limited means to a cardiac consultant for complete diagnostic study.

Since there is implicit in a screening program acceptance of responsibility for identification

![Figure 2](image-url)

**Fig. 2.** The square green and red figures previously selected by the screen (fig. 1) are brought back for diagnosis (the assistance of the public health nurse is obvious). A thorough diagnostic work-up should separate the positives, the square red figures, from the false positives, the square green figures, refer the former to medical treatment while the latter are sent on their way. *(The use of color in this illustration is made possible by a grant from Wyeth. Incorporated to the publication fund of the American Heart Association.)*
and is of no real additional help in establishing a diagnosis. Until an effective secondary screen is devised, complete diagnostic facilities, particularly for those with limited means, may be necessary to provide the patient’s physician with adequate information and to assure a successful survey.

**Principles Governing Use of Screening Tests in Mass Surveys**

When a screening test has been deemed satisfactory, its effectiveness in a mass survey can be evaluated by:

(1) **Yield**

Many methods of measurement of the yield of a screening program have been proposed, including:

(a) **The number of previously unknown verified cases of disease among the total population surveyed.** This index places emphasis on lack of previous knowledge of disease, but makes no reference to any benefit which would accrue to the discovered cases. In the first of the Boston studies, where trained and experienced radiologists read the films, there was a yield among the surveyed population of 0.9 per cent of verified cases of heart disease previously unknown.\(^1\)

(b) **The number of persons with previously unknown verified disease benefited by referral to medical care, and the number of previously known cases not under medical care, benefited by return to it.** In the first of the Boston studies, slightly less than half of the verified cases of heart disease previously unknown were believed to have been benefited,\(^1\) but no data were available on the number of previously known cases of verified heart disease who were returned to medical care.

(c) **The number of individuals who believe they have the disease, have a positive screening test, but are shown not to have the disease by subsequent diagnostic examination.** This yield may be considered additive to the yield as measured under (a) or (b) above. In the Boston studies, there were 18 individuals in group I and nine in group II who fell into this category.\(^1\)

(d) **The number of cases of communicable disease who are prevented from spreading their disease to the family or the community.** Heart disease is not infectious and therefore no yield can be calculated in this category.

In considering these various indexes of yield in the Boston studies, it is believed that the yield of previously unknown cases discovered and benefited in the group I study would be considered satisfactory, in comparison with yields usually obtained in mass surveys for other diseases.

(2) **Cost**

The yield of a screening program must be balanced against the cost. Cost is measured in monetary terms, which are in turn affected by the relative use of the time of professional and nonprofessional personnel.

It is impossible to calculate the cost of the Boston heart disease studies since they were organized as studies and not as routine case finding mass surveys. The monetary costs were not limited to the additional procedure of reading for heart disease the films taken for tuberculosis, the cost of the follow-up and diagnostic facilities and procedures for referral for care, but included also the cost of facilities and personnel responsible for collection and analysis of the data.

When one attempts to measure monetary costs of the Boston studies exclusive of the research costs, they are very variable depending upon which items are to be included in the calculations. Thus in these studies they included not only the basic cost of the tuberculosis survey and the additional charges imposed by heart disease case finding, but also many hidden costs in permanent personnel and facilities in the agencies responsible for the organization and operation of the studies and the amount contributed in free time by professional and lay groups. However, each community can calculate in advance the approximate additional cost of reading 70 mm. films for heart disease, and the cost of follow-up, diagnosis and referral for care, based on the particular procedures to be used in the community in question.*

* The special readers in the group I study estimated that in the average community survey for
The practical limit in the number of professional personnel, particularly physicians with highly specialized training, demands that screening tests and programs use a minimum of professional time and a maximum of the time of qualified technicians and other non-professional personnel. Moreover, the monetary cost will increase pari passu with the increased use of professional personnel. This statement may be paraphrased as "Never use an expert as a screen." This statement relates only to community-wide surveys and not to the very useful screening function provided by the general practitioner in recognizing disease in its early stages in patients who come to his office with early symptoms.

The screening procedure which has been most successful, that is, the 70 mm. photofluorographic film for tuberculosis, is based on this principle. In contrast, the cancer detection center demands much time from highly specialized physicians, the cost per case found is high, and this method of cancer screening has consequently been unsuccessful. In mass surveys the time of highly trained physicians is more efficiently used for precise diagnosis of disease in those already selected by the screen.

The method of heart disease case finding using the 70 mm. photofluorographic film can be performed by average technicians. Although the films require interpretation by a physician trained in the reading of such films for heart disease, the amount of "expert time" is relatively small. Thus this survey method is a practical one.

(3) Acceptance*

Reliability, validity, yield and cost are essential criteria for the evaluation of screening tests and programs. The measurement of acceptance of the program by individuals, physicians and the community is a useful additional criterion of effectiveness of a screening program.

Acceptance of the program by the individual is measured by the proportion of the population which offers itself for screening, and then by the proportion which accepts the results of the test and follows recommendations. The patient's acceptance of the program is probably dependent, among other things, on the motivation of the individual who comes or who does not come to the screening program, and the emotional reaction to, and pain, embarrassment or hazards involved in, the testing procedures. Very little quantitative information is available on these points. Their measurement requires that scientific techniques employed in the study of social attitudes and behavior be applied to the study of this problem.

Studies of the effect of mass screening programs on the case load imposed on the physician should be useful in measuring physician acceptance, which is a sine qua non of the success of the screening program.

Acceptance by the community can be measured by the degree of responsibility taken by official and voluntary community agencies to increase funds and facilities for care of patients discovered by the screening program, and by the spirit with which the community undertakes subsequent screening programs.

In the Boston heart disease studies, professional groups including physicians, nurses and many community agencies concerned with heart disease cooperated effectively to assure the success of the program. In these studies, patient acceptance is evidenced by the return for diagnosis of over three-fourths of those whose films were read as abnormal for size and shape of cardiac silhouette; however, this high rate of return for diagnosis required concerted follow-up efforts of family physicians, the Boston Visiting Nurse Association and the study group.

Special Population Groups

Although mass screening tests and programs are designed for case finding in the general population, they may be applied to special population groups: school children, industrial workers, social and racial groups and others,
when there is evidence that a high yield may be expected.

For example, in the Boston heart disease studies, higher yields were obtained in females than in males and in nonwhites than in whites. It might be expected therefore that in surveys of females or of nonwhite populations high yields would be obtained, and the yield should be highest of all in a special population group of nonwhite females.

**Multiple Screening Programs**

Since the first demonstration of a successful multiple screening program it has been recommended that the simultaneous application of many mass screening tests is more efficient than many individual mass screening programs. The evaluation of a multiple screening program is in principle no different from that of a screening program for a single disease. Each of the component tests and programs must be evaluated in accordance with the above criteria. In addition, special studies are needed to determine whether a particular multiple screening program increases or decreases the effectiveness of the individual component tests, and whether it is more effective than a series of individual screening programs. When these conditions have been satisfied, it may then be possible to recommend the combination of individually effective mass screening procedures into a multiple screening program.

In the Boston studies, the addition of case finding for heart disease to the tuberculosis case finding program in a sense created a multiple screening procedure. The results attained in group I justify the use of the 70 mm. photofluorographic tuberculosis survey film as a heart disease case finding procedure, whether performed individually or jointly with other established mass screening tests and programs.

* The first paragraph of this section, with minor changes in wording and with the addition of the last sentence, is adapted from the report of the Committee on Early Detection and Screening, Section on Evaluation of Scientific Data, National Conference on Chronic Disease: Preventive Aspects.

**Acknowledgments**

The author is indebted to Mr. Felix E. Moore of the National Heart Institute, United States Public Health Service, and to Miss Marjorie T. Bellows and Mrs. Ruth E. Lynch of the American Heart Association, for reviewing the manuscript of this paper and for their excellent suggestions and advice.

**References**
