MULTIPLE SCREENING IN BALTIMORE

THESIS

Submitted to the Faculty of Medicine of the University of Glasgow in conformity with the requirements for the degree of Doctor of Medicine.

Charles Murray Wylie

August 1956

ProQuest Number: 13848976

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 13848976

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENT

I would like to acknowledge my great indebtedness to the staff of the Commission on Chronic Illness of the United States of America for the help and advice freely given to me while carrying out this work. I am particularly indebted to:

Dr. Dean W. Roberts, Director.

Mr. Dean E. Krueger, Research Analyst.

Mr. Peter G. Meek, Assistant Director.

Mr. Marvin Glasser, Research Assistant.

I also acknowledge with gratitude the kindness shown by Dr. Ernest L. Stebbins, Professor of Public Health Administration, Johns Hopkins School of Hygiene and Public Health, under whose supervision this work was carried out.

Charles M. Wylie.

CONTENTS

SECTION A

GENERAL ASPECTS OF MULTIPLE SCREENING

Chapter 1

Introduction

INTRODUCTION

A vigorous controversy exists over the present use and future possibilities of multiple screening. This controversy is well shown by the following statements made by two public health experts:

"The concept of multiple screening," says Dr. L.A.Scheele, Surgeon General of the United States Public Health Service, "is basically sound and fits in admirably with the current trend in public health work generally, toward a unity and greater integration of programs than we have had in the past." (Scheele, 1951.)

Opposing this point of view, Professor Wilson G. Smillie (1952) states: "But a multiphasic diagnostic screening unit that is established as a primary public health diagnostic function is unsound in its concepts, untenable in its principles and indefensible in its logic."

Much information exists on multiple screening, but is widely scattered in the literature. The secondary purpose of this thesis is to bring together much of this information to help clarify the present confused situation.

The primary purpose of this thesis is to present the results of a multiple screening clinic held in Baltimore, Maryland, U.S.A., during the fourth October to the twelfth December, 1954. This clinic was one step in a study of the prevalence of chronic illness in Baltimore, which was being carried out by the Commission on Chronic Illness.

Definitions.

Screening. In its public health sense, screening is a method which sorts out persons who probably have abnormalities from those who probably do not (Adapted from Commission on Chronic Illness, 1951.)

Screening tests are procedures which sort out persons who probably have abnormalities from those who probably do not (Adapted from Commission on Chronic Illness, 1951.)

Screening tests are applied to groups of apparently well persons. The tests are procedures preliminary to referring persons with positive results for diagnosis and treatment if necessary. Screening tests do not establish or rule out disease; diagnostic procedures perform that function. Chapter 2 of this section will discuss more thoroughly the difference between screening and diagnostic tests.

History of Screening.

Screening tests have been used in public health practice for many years. For instance, the Rockefeller Sanitary Commission used the microscopic examination of stools to screen for hookworm disease in and after 1912 (Smillie, 1952.). The miniature chest x-ray for pulmonary tuberculosis has been probably the most widely used of all screening tests. First used about 1930 in the United States and about 1936 in Great Britain, the miniature chest x-ray continues to play an important part in tuberculosis control.

Multiple screening is a natural and perhaps inevitable outgrowth from single screening programs. Multiple Screening is the rapid and economical use of two or more screening tests to refer for diagnosis and treatment, if necessary, those persons who probably have abnormalities (Adapted from Commission on Chronic Illness, 1951.) . Such programs have also been named multiphasic screening, mass screening or multi-test programs.

Early American examples of these programs tested special groups, such as armed forces entrants and industrial workers. For instance, a combined syphilis and tuberculosis screening program tested industrial workers in Richmond, Virginia in 1939. Multiple screening programs for the general population came later, such as the Georgia State Health Department's program beginning in 1945.

The first brief but intensive multiple screening survey was held in San Jose, California in 1948 (Canelo et al.,1949.) This program, limited to industrial workers, was followed by more extensive programs open to the general public. Among the best known was the Richmond, Virginia multi-test clinic which tested some 38,000 persons in 1950 (Boek, 1951).

Despite the advantage that combining several individual screening programs will save money and effort, multiple screening has not reduced the growth of single testing programs in the United States. The miniature chest x-ray programs, for

instance, continue on a large scale. Diabetes Detection Drives have become a nationwide annual event since 1948, following the successful United States Public Health Service Survey in Oxford, Massachusetts, in 1946 (Wilkerson and Krall, 1947).

The years 1953 to 1955 have seen fewer reports of multiple screening programs in America than previous years. There is now a tendency to study more thoroughly the previous programs, to determine whether they have been worth while. Smaller numbers of persons have been tested, to leave energy and funds for the study of what happens to positive screenees after their referral.

In Great Britain the miniature chest x-ray continues to be the most widely used screening test. School clinics use audiometric and visual acuity tests, while tests for syphilis, hypertension and albuminuria are much used in antenatal clinics. Multiple screening programs in the United Kingdom have involved only small numbers and select groups of the general population. Since 1949, for instance, the Mass X-ray Service for South West London has measured heights, weights and erythrocyte sedimentation rates in addition to taking the chest x-rays (Nash, 1952). Other programs have involved corporation staff volunteers (Burn, 1952) and university students (Logan, 1954). Recently several British workers (Nash, 1952; Anonymous, 1955) have recommended that larger programs be carried out to define the part which multiple screening might play in the National Health Service in future years.

REFERENCES

Anonymous,(1955). Medical Officer XCIV,339.

Boek,W.E. (1951). An Analysis of the Multi-test Clinic of Richmond, Virginia. Health Information Foundation, New York.

Burn, J.L. (1952). Lancet 262,567.

Commission on Chronic Illness (1951). Preventive Aspects of Chronic Disease, Conference Proceedings, Chicago, Ill. Health Publications Institute, Inc., Raleigh, N.C.

Canelo,C.K.,Bissell,D.M.,Abrams,H., and Breslow,L.A. (1949) California Med. 71,409.

Dunn,J.E. (1955). J.Chronic Dis. 2,450.

Logan,R.F.L. (1954). Lancet 2,409.

Nash,F.A. (1952). Lancet 1,673.

Scheele,L.A. (1951). Am.J.Pub.Health 41,635.

Smillie,W.G. (1952). Am.J.Pub.Health 42,255.

Wilkerson,H.L.C., and Krall,L.P. (1947). J.Amer.med.Ass. 135,209.

PHILOSOPHY AND PRACTICE OF SCREENING

Section A
Chapter 2

PHILOSOPHY AND PRACTICE OF SCREENING

With aging of the population and the falling incidence of infectious disease, chronic illness becomes a steadily increasing medical and public health problem. Since much chronic disease is of uncertain etiology, its primary prevention is not possible at the present time.

Secondary prevention -- the earliest possible detection of already existing disease (Levin and Brightman, 1952) -- may be the most practical alternative in controlling chronic illness. The widespread use of one method of secondary prevention -- the periodic health examination -- has been impractical for these reasons:

1. The number of physicians is limited, and most are already fully occupied in caring for ill patients.
2. Undergraduate and postgraduate medical training lays little stress on early diagnosis and the preservation of health.
3. The average layman usually seeks medical attention only when ill.
4. A good periodic health examination is costly to the individual if carried out privately, or to society if financed by public funds.

Periodic health examinations may be more feasible if confined to persons most likely to have conditions which benefit from medical care. Multiple screening tests have thus been proposed to separate the population into (a) persons who probably have conditions requiring medical supervision, and (b) those who probably do not have such conditions.

Before developing this theme, one might briefly consider the four basic points in this line of reasoning:

1. The most effective method of control of any chronic disease is its primary prevention before the disease begins. Physicians appear to agree unanimously in this belief, because of the success of primary prevention in communicable disease control.

2. Primary prevention of chronic illness is rarely possible, since the etiology of most chronic illness is not clear. Treating chronic disease at its earliest detectable stage is therefore the most practical alternative. Again the medical profession will agree almost unanimously on this statement. The more advanced the disease is when treatment begins, the more difficult is its complete cure. However, carrying out this step is again greatly hindered by insufficient knowledge of many of the chronic diseases. Often their treatment is symptomatic in nature, aimed at relieving the discomfort produced by the condition. Symptomatic treatment will not help when illness is discovered before the symptoms develop. Moreover the assumption has still to be proved that asymptomatic disease is more likely to be pathologically 'early' disease.

3. The third basic point in this reasoning is that periodic medical examination of apparently well persons will discover asymptomatic disease at a stage when relief and even cure is more simple. The medical profession does not agree on this point. Indeed, a recent editorial (1954) put forward what is presumably an official view of the British Medical Association,

that periodic examination of apparently well persons has little value.

Periodic visits to the doctor will not ensure good health, and may even occasionally cause harm (Bakwin, 1945). The periodic health examination is merely a means to the goal of early treatment, and is not an end in itself. Poor and hastily administered examinations will prevent the attainment of this goal, as will the overenthusiastic "treatment of normal variants" (Bakwin, 1945) which occasionally results from these examinations. In brief, the periodic examination needs evaluation to find out if it prevents enough illness and saves sufficient lives to justify its cost.

4. The fourth basic point in reasoning the need for screening tests is that the tests will separate off persons who will benefit most from a thorough examination. There is no agreement that such screening procedures are available. Indeed, there is general agreement that no screening tests are presently available for many chronic conditions. It is claimed, however, that some of the most prevalent conditions, which are capable of responding to medical treatment, can already be screened successfully. Much of this thesis will be devoted to studying this claim.

Comparison of Screening and Diagnostic Tests.

The differences between screening and diagnostic tests might suitably be developed at this point.

Screening tests sort out persons who probably have abnormalities from those who probably do not. Screening tests must

be rapid, inexpensive and simple to administer. In order to attract apparently well persons to attend, screening tests should be easily acceptable or accompanied by a minimum of discomfort. Painful or unpleasant tests discourage the attendance of large numbers of persons.

Screening test results should be simple to interpret, so that the need for referral may be decided, if possible, by non medical personel. Since a positive or negative result is all that is necessary, the tests need not give specific values in the result. A screening test for anaemia, for instance, may not give the actual haemoglobin value in gm. per cent, but need only show that the result is above or below the value where further examination is advisable.

Inherent in the philosophy of screening is that screening tests should detect conditions which will benefit from medical care. Detecting conditions for which no effective treatment is known has little place in screening, except where prevalence figures are required or where isolation of the cases has some public health value.

Diagnostic tests are usually applied to apparently sick persons or to persons with positive screening tests. Diagnostic tests help the physician to make a definite diagnosis. Since they are applied on an individual basis, diagnostic tests are often less rapid, more expensive and more complicated than screening tests.

Since diagnostic tests are given to apparently sick persons who are anxious for the diagnosis to be made, the production of discomfort is not a barrier to their use. Diagnostic tests usually give specific values as their results, which are often more difficult to interpret than those of screening tests.

Diagnostic tests do not always detect illness for which there is adequate treatment. An accurate diagnosis is often the physician's immediate concern, and the absence of effective treatment is no barrier to using the diagnostic test.

Results of Screening.

Screening tests usually give a positive or negative result, indicating that the screenee should or should not be referred for further examination. However, screening tests cannot do this job with complete efficiency. The tests will wrongly refer some persons who do not have the conditions at which the tests are aimed. The tests also will wrongly classify as not requiring referral some screenees who actually have the conditions. The following terms are used to describe the results of screening:

True positives are the results of persons selected for referral on screening and found positive on diagnosis.

False positives are the results of persons selected for referral and found negative on diagnosis.

True negatives are the results of persons not selected for referral on screening and found negative on diagnosis.

False negatives are the results of persons not selected for referral on screening and found positive on diagnosis.

Figures for the latter two groups are not usually available in practice, since the negative screenees seldom have a diagnostic examination.

Evaluating Screening Tests.

Several different aspects of screening test performance have been recommended to evaluate the tests (Levin and Brightman, 1952). These aspects are:

1. Reliability. The reliability of the test can be measured by
   a. Its precision, or ability to give closely similar results when repeated on the same person.
   b. Its accuracy, or ability to give results close to the actual true value.
2. Validity. The validity of a test is its ability to separate the true positives from the true negatives, keeping to a minimum the number of false positives and false negatives. Two indices are used to express validity:

a. The sensitivity of a test is its ability to classify as positive persons who have the condition at which the test is aimed. Numerically, this may be written as

$$\text{Sensitivity} = \frac{\text{Positive screenees, confirmed pos. on diagnosis}}{\text{Total screenees who have condition}}$$

$$= \frac{\text{True positives}}{\text{True positives + False negatives}}$$

b. The specificity of a test is its ability to classify as negative persons who do not have the condition at which the test is aimed. Numerically, this may be written as

$$\text{Specificity} = \frac{\text{Negative screenees, confirmed neg. on diagnosis}}{\text{Total screenees who do not have condition}}$$

$$= \frac{\text{True negatives}}{\text{True negatives + False positives}}$$

3. Yield. Unlike the previous indices of performance, there is no general agreement on a definition of yield. Rutstein (1951) suggests several ways in which the yield can be expressed numerically, including:

a. Number of previously unknown verified cases.

b. Number of previously unknown verified cases benefited by referral, and previously known cases not under treatment, benefited by return to it.

c. Number who believe they have the disease, have a positive screening test, and who are shown not to have the disease on diagnostic examination. (While those persons do benefit from screening, it does not seem rational to count this in favour of a test which has wrongly classified these persons as positive.)

d. Number of cases of communicable disease prevented from spreading the condition. (These persons will, of course, be included in group a, when the latter is used to express the yield.)

Rutstein (1951) emphasises that only persons who benefit from referral should be included in estimating yield. Although excellent in theory, this idea is difficult to carry out in practice, since it will considerably increase the period elapsing before follow-up figures are obtained. The procedure will require time to determine the response to treatment. Moreover the procedure involves a purely subjective decision on the

part of the personal physician. Some physicians may believe that all patients referred to them will benefit by the referral, while others may take a less optimistic view of their therapeutic powers. The decision may vary greatly from physician to physician and from area to area. Thus the same test producing the same proportion of true positives in two different areas might appear to have quite different yields.

The problem of delay in obtaining the final follow-up figures will be less serious where the screening program is a continuing one over the course of one or more years. However, the uniform interpretation of "benefiting" from the treatment will continue to be a problem.

For the sake of simplicity and comparability of results, therefore, the most satisfactory index of yield appears to be group a -- the number of previously unknown positives confirmed by diagnosis.

A practical question must be answered -- to whom should the condition be "previously unknown"? The screenee, for instance, may be unaware of his heart condition because his personal physician believes it wise to leave him in ignorance. It therefore seems rational that the condition be "previously unknown" to the personal physician, and not to the screenee. This decision has a considerable effect on the yield of the test, as the results of the Baltimore clinic will show.

The test yield is a function of the prevalence of unknown cases of the condition at which the test is aimed -- the greater the prevalence, the higher the yield. When the prevalence does

not greatly vary from area to area, the yield is helpful in evaluating the performance of two different tests aimed at the same condition. The yield is even more helpful when the two tests are used on the same population. On the other hand, the yields of two tests aimed at two different conditions with different prevalences will not evaluate the test performance alone.

4. Cost is the fourth aspect to be considered in evaluating screening tests. While screening studies sometimes give the cost per person tested, the cost per new confirmed case gives a truer cost picture of particular value to the screening clinic administrator. Of importance to the community, however, is the total cost, including estimates of volunteer services and donated materials, and estimates of the cost of diagnosing the true positives and false negatives, all added and expressed in terms of new cases found. Optimistic accounts of screening costs have been calculated by ignoring the hidden costs to the community, which are often very considerable.

A suggested fifth aspect (Rutstein, 1951) -- the acceptance of the screening program by the screenees, by the physicians and by the community -- is so dependent on factors apart from the screening tests that this method seems quite useless for evaluating screening tests.

REFERENCES

Bakwin, H. (1945). New England J.Med. 232, 691.

Editorial (1954). British Med. J. 2, 1095 (Nov.6).

Levin, M.L.,and Brightman, I.J. (1952). New York State J.Med. 52, 2600.

Rutstein, D.D. (1951). Circulation IV,659.

General References.

Breslow,L., and Roberts,D.W. (1955). Introductory Statement on Screening for Asymptomatic Disease, J.Chronic Dis. 2, 363.

Commission on Chronic Illness (1951). Preventive Aspects of Chronic Disease, Conference Proceedings, Chicago, Ill. Health Publications Institute, Inc., Raleigh, N.C.

Levin,M.L. (1955). Screening for Asymptomatic Disease, J.Chronic Dis. 2,363.

PERFORMANCE INDICES FOR SCREENING TESTS

Section A
Chapter 3

PERFORMANCE INDICES FOR SCREENING TESTS

A single index to describe test performance would greatly help in investigating screening tests. A satisfactory index should change only with changes in the test, and be unaffected by changes in prevalence of the condition being tested for in the population. To discuss the value of several proposed indices, the results of a theoretical screening test are tabulated in the following table:

		Classified by SCREENING		
		Positive	Negative	Total
Classified by DIAGNOSIS	Diseased	a True +	b False -	a + b
	Healthy	c False +	d True -	c + d

An index already described can now be expressed as:

$$\text{Sensitivity} = \frac{a}{a + b}$$

The greater the sensitivity, the larger the proportion of diseased persons correctly classified as positive. When test results are compared from two different areas, the numerator and denominator of this index are both directly proportional to the number of diseased persons in the tested population. Whether the prevalence of the condition being tested for is one or ten per cent, both numerator and denominator will increase by the same proportion and the index will not change its value.

The index will change, however, when an artificial population is tested where all diseased cases are far advanced.

This situation may occur, for instance, if a new cancer detection test were evaluated using a group of hospital inpatients with cancer and a "normal" group of apparently healthy medical students or nurses. When used on this artificial population, the test will have a sensitivity different in value from the result obtained when the test is used on the general population, where the cancer cases will be at an earlier stage. The use of artificial populations will affect most other indices used to describe the performance of screening tests.

Consider an entirely imaginary and useless test which records all results as positive. The index of sensitivity will be a/a or 100 per cent, since no false negatives have occurred. This extreme example illustrates the fact that although the sensitivity is a useful index, it will not give a complete picture of the performance of the test.

The other index mentioned previously can now be expressed as:

$$\text{Specificity} = \frac{d}{c + d}$$

The less specific the test, the larger will be the proportion (c) of healthy persons wrongly classified as positive. A low specificity indicates that an excessive strain may be placed on the limited facilities available for diagnosis and treatment.

Since both numerator and denominator are directly proportional to the number of healthy persons in the tested population, the specificity does not alter with changes in prevalence of the condition. Once again, however, the specificity may change if the test were used on artificial populations.

In the extreme case mentioned previously of the test which classified all persons as positive, the specificity would be 0 per cent. In the opposite extreme of a test classifying all persons as negative, the specificity would be 100 per cent while the sensitivity would be 0 per cent. Both sensitivity and specificity are useful indices when considered together, but neither gives a complete picture of test performance when only one is known.

Clinicians object to tests with a significant number of false negatives, since such persons may have treatment delayed until the disease is far advanced. On the other hand, health administrators, and screenees classified wrongly by the screening process, tend to react adversely to tests resulting in numbers of false positives. Such tests tend to cause considerable unnecessary worry and expense, and may overburden the diagnostic and treatment facilities.

An index which stresses the latter situation is the ratio of False positives:True positives or c/a. This index has considerable value in showing the amount of unnecessary follow-up work produced by the test.

However, the false positive:true positive ratio will change when the same test is used on groups with differing prevalence of the condition being screened. The number of false positives, which is the numerator of this index, is directly proportional to the number of healthy persons tested. The true positives, which form the denominator of the index, are directly proportional to the number of diseased persons tested.

A high proportion of diseased persons in the tested group will reduce this index, even though the efficiency of the test is unchanged. Provided this defect is remembered, however, this ratio has some value in evaluating the performance and practicability of a test.

When used alone, each index considered so far has clearly failed to describe completely the performance of a test. Two indices have been suggested to give a more complete picture of test performance.

The numerator of the first of these indices is the number of persons correctly classified by the test -- the true positives and the true negatives. The denominator is the total number of persons tested -- the true and false positives and negatives. The index can therefore be written as

$$\frac{a + d}{a + b + c + d}$$

An insensitive and non-specific test will give a low value, while a sensitive and specific test will give a high value approaching unity or 100 per cent.

However, this ratio has a defect that almost precludes its use completely. Take the extreme example of a test giving a negative result for everyone, in a population with a one per cent prevalence of the disease. This situation would give a value of 0.99 or 99 per cent for this index. The basic defect is that this ratio gives equal weight to the correct classification of a diseased or a well person. Where the disease prevalence is low, this index can give a false appearance of efficiency to a test which, in actual fact, is completely useless.

Youden (1950) proposed and index, designated by the symbol "J", which appeared to overcome the previous defect. Youden's reasoning was as follows:

Returning to the fourfold table, the proportion of diseased individuals correctly classified is a/(a + b). It seems appropriate to charge against thistthe proportion of diseased individuals wrongly classified, namely b/(a + b). The measure of success on the diseased group will thus be $\frac{a - b}{a + b}$. If the test has no discriminative power on the diseased group and is equally likely to report a diseased individual negative as positive, a and b will be equal and the numerator of this fraction will be zero. The test with no false negatives will have b = 0 and $\frac{a - b}{a + b}$ equal to one.

A similar argument for the healthy group gives $\frac{d - c}{c + d}$ as a measure of the effect on the well persons. Let the average of these two fractions be taken as the index of performance of the test on the whole population. This gives equal weight to the wrong classification of a certain percentage of diseased persons as to the wrong classification of the same percentage of healthy persons. This method is better than the previous index, although it still has serious defects which will be illustrated later.

We now have $J = \frac{1}{2}\left(\frac{a - b}{a + b} + \frac{d - c}{c + d}\right)$

Within the parentheses, add one to each fraction, and subtract 2.

Then $$J = \frac{1}{2}\left(\frac{2a}{a + b} + \frac{2d}{c + d} - 2\right)$$

$$= \frac{a}{a + b} + \frac{d}{c + d} - 1$$

$$= \text{Sensitivity} + \text{Specificity} - 1$$

The suggested index is simply obtained, therefore, by adding the sensitivity and specificity, expressed as fractions, and subtracting one. It is claimed to have certain desirable features, which are more thoroughly explained in Youden's (1950) original paper:

1. When its value is zero, the test has completely failed to differentiate between diseased and healthy persons. (This also points out the fact that when Sensitivity and Specificity add up to one, the test gives the same proportion of positive results among diseased persons as among healthy persons. Their sum can be less than one only when healthy persons give a higher proportion of positive results than diseased persons i.e. when a negative result suggests the presence of disease.)
2. When its value is one, the test produces no false positives and no false negatives.
3. When its value is between zero and one, this value is controlled by the proportion of diseased persons wrongly classified as negative, and by the proportion of well persons wrongly classified as positive.
4. The index is independent of the prevalence of the disease in the population and of the total number of persons tested.
5. It is possible to calculate a standard error for the index, and thus to evaluate the difference in performance of two tests.

An actual example will be given, however, to illustrate a serious defect of the index. Consider the results of classifying 1000 well persons and 10 ill persons, using Test A with a sensitivity of 0.9 and specificity of 0.7. This test will

produce one false negative and 300 false positives. 301 persons are wrongly classified, and 309 are referred for diagnostic examination to discover 9 true positives.

Test B, applied to the same population, has a sensitivity of 0.8 and specificity of 0.8. 202 persons are wrongly classified, and 208 are referred for diagnostic examination to discover 8 true positives.

Although tests A and B have the same J of 0.6, their end results are quite different. It therefore appears that the Youden Index is not sufficiently informative to compare the efficiency of two tests. However, even if it has only emphasised the fact that a test is completely useless when its sensitivity and specificity add up to one, the index has performed a helpful function.

When A test with a low specificity, giving a fairly high proportion of false positives, produces serious practical difficulties in requiring much unnecessary follow-up. It would seem rational to fix some high specificity requirement for the screening tests. Their indices of sensitivity might then be compared, keeping their specificities at this constant level. An example will be given to illustrate this procedure.

Suppose that a urine sugar test and blood sugar test are carried out on 1000 well persons and on 15 diabetic persons. It is arbitrarily decided that the specificity will be fixed at 0.96, thus permitting 40 false positives. It is found, let us say, that 46 persons have a positive urine sugar test, and that 40 of these are non diabetics and 6 are diabetics. These

results give the required specificity of 0.96, and a sensitivity of 6/15 or 0.40.

The blood sugar test has, say, 50 positives, with 40 non diabetics and 10 diabetics. Again these results give the required specificity of 0.96, with a sensitivity of 10/15 or 0.67. These hypothetical figures suggest that the blood sugar test, with the higher sensitivity for the same specificity, is the more efficient screening test.

This method of comparison can be used only when the total number of well and diseased persons are known, and when the screening levels can be altered to give the required specificity. For instance, the actual blood sugar values of the screenees must be available, so that the screening level can be changed from one point to another after the tests are completed.

In the example just given, the two tests detect the same condition. This method of comparing two tests is less useful when the tests screen for different diseases.

The specificity can be arbitrarily fixed at any value to compare the sensitivities. However, the most reasonable value to choose would appear to be the "minimum permissable specificity". This value will partly depend on:

1. The possible damage -- psychological, financial, etc -- caused by the wrong classification of the truly negative person.
2. The seriousness of missing a case of the condition at which the test is aimed. For instance, when the missed case will

progress rapidly to an incurable stage, the specificity may deliberately be permitted to be low in order to reduce the number of missed cases.

3. The prevalence of the condition in the tested population. When the prevalence is relatively high, such as with "hypertension" or "heart disease", a comparatively low specificity may be used if desired since the false positive:true positive ratio will not become unduly high.
4. The response of the condition to medical supervision. If only the relatively "advanced" cases require medical care (as may be the case with essential hypertension) the specificity may be held at a high value in order to avoid the unnecessary referral of the "early" cases.

Conclusion and Summary.

No single index will give a complete picture of screening test performance. However, a screening test can be evaluated adequately by an intelligent study of the indices of sensitivity and specificity, and of the ratio of false positives: true positives.

REFERENCE

Youden,W.J. (1950). Cancer 3, 32.

General References.

Breslow,L., and Roberts,D.W. (1955). Introductory Statement on Screening for Asymptomatic Disease, J.Chronic Dis. 2, 363.

Dunn,J.E., and Greenhouse,S.W. (1950). Cancer Diagnostic Tests. Principles and Criteria for Development and Evaluation. Public Health Service Publication No. 9. U.S.Government Printing Office, Washington, D.C.

Greenhouse,S.W., Cornfield,J., and Homburger,F. (1950) The Youden Index: Letters to the Editor, Cancer 3, 1039.

Levin,M.L. (1955). Screening for Asymptomatic Disease, J.Chronic Dis. 2,367.

STATISTICAL ASPECTS OF SCREENING

Section A
Chapter 4

STATISTICAL ASPECTS OF SCREENING

When a population is screened for a condition such as diabetes, it is presumed that the diabetic will give a different test result from the non diabetic. However, the range of results obtained from the diabetics will considerably overlap the range obtained from non diabetic persons.

In order to simplify the screening procedure, a point is arbitrarily chosen somewhere on the overlapping portion. On one side of this point lie the "positives", to be referred for more thorough examination; on the other side lie the "negatives!" This point has been variously termed the critical point, cutting point, screening level, critical level and cutting level.

The population to be screened at this level consists, let us say, of the two groups shown in Figure I. The normals have the distribution of values at the lower range of the scale, while the diabetics have the values distributed over the upper range.

Suppose that this is a group of 10,000 non diabetic persons and 150 previously unknown diabetics, having blood sugar tests which are being carried out without considering the history of food intake. Non diabetic persons who have recently eaten a large meal will often have blood sugar values above those diabetics who have not eaten for several hours. There will therefore be a considerable overlap of the two distribution curves as suggested in Figure I. No matter what screening level is used, a considerable proportion of false negatives or false

Theoretical Distributions of Blood Sugar Values of 10,000 Non-Diabetic Persons and of 150 Diabetics.

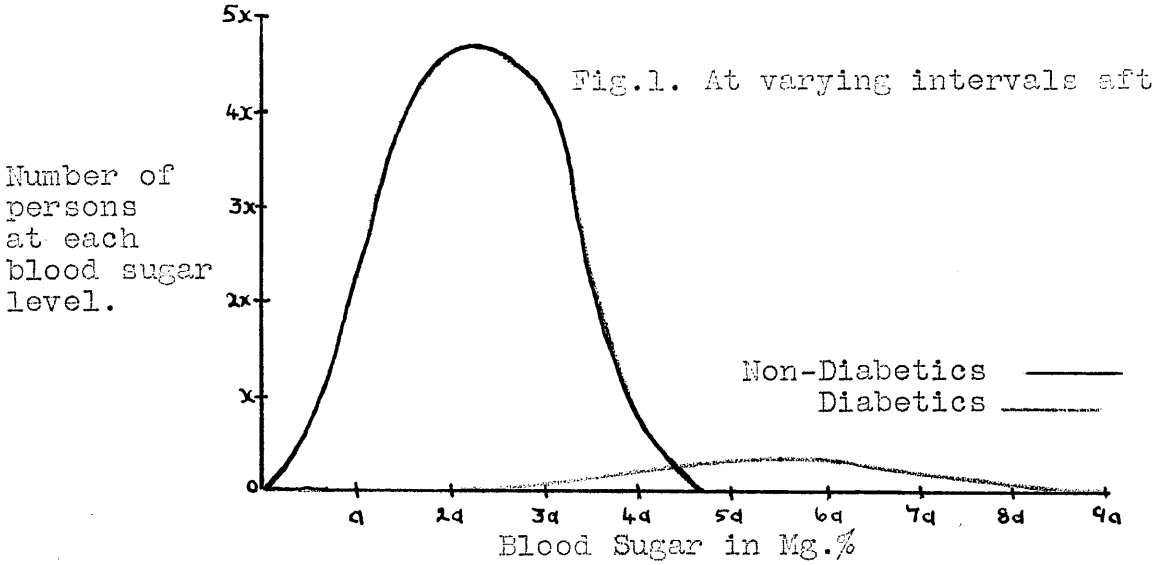


Fig.1. At varying intervals after food

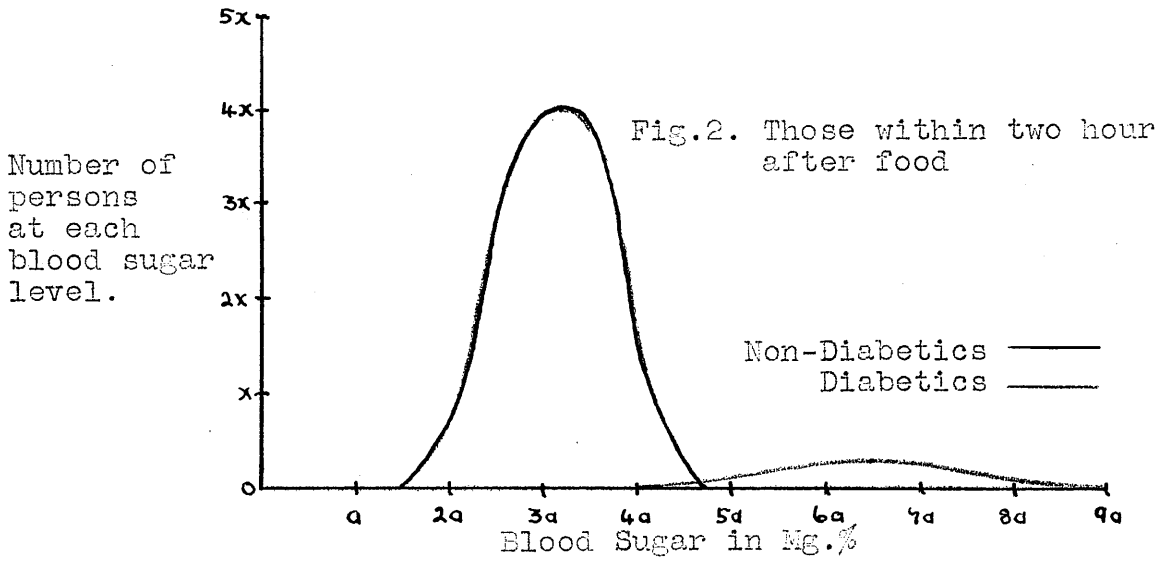


Fig.2. Those within two hours after food

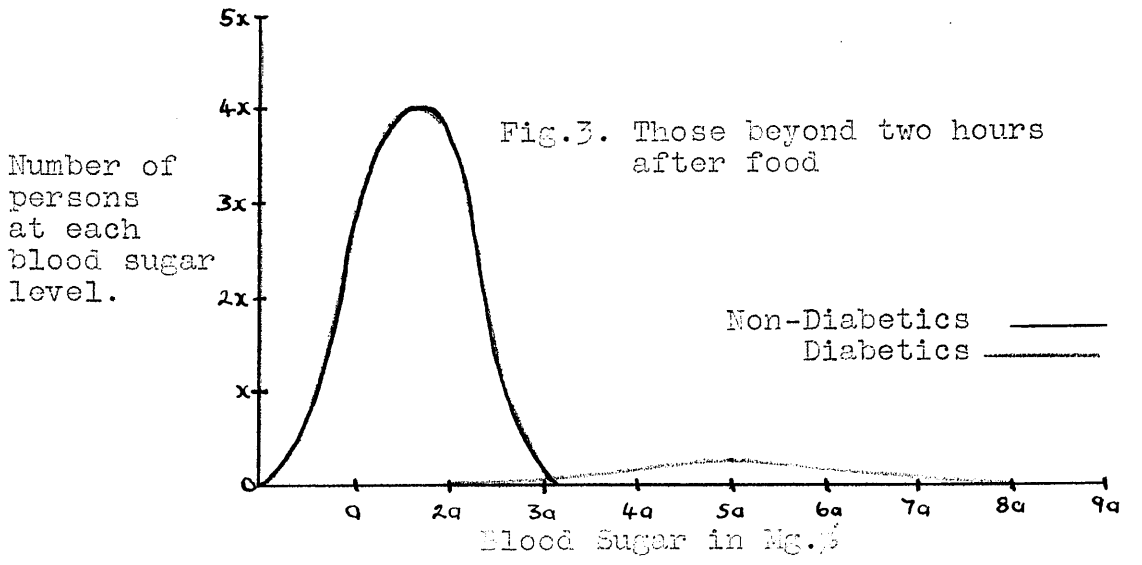


Fig.3. Those beyond two hours after food

positives or both is inevitable. (There is no evidence, incidentally, that the two distributions will be "normal" in shape, although they are shown so here.)

Reducing the overlap of the two distribution curves would make the screening process more efficient. This reduction occurs if the means are placed further apart, without increasing the distribution spread; if the means remain stationary but the distribution curves diminish; or if a combination of both situations occur. In the case of diabetes screening, the overlap can be reduced by several methods, all of which can be described as "standardising the conditions under which the test is carried out."

The population screened can be divided into two groups: those who have eaten within two hours, and are still likely to be absorbing some food from the gastro-intestinal tract; and those who have eaten beyond two hours, who have mainly completed the process of absorption. Using a higher screening level for the first group, say 160 mg. per cent, and perhaps 130 mg. per cent for the second group, will result in a smaller overlap of the two curves as illustrated in Figures 2 and 3.

Another method of reducing the overlap is to give each person a glucose drink at a specific time before the test. This procedure will slightly raise the mean value for the "normals" and considerably raise the mean for the diabetics, without greatly increasing their standard deviations.

The Test Error.

Suppose that 500 ml. of blood, taken from one "normal" person, were divided into 500 lots of one ml. and handed to a technician to estimate the blood sugar level on each specimen. The technician has not been informed that the blood comes only from one person with a blood sugar value of, say, 100 mg. per cent. If the test used has an equal chance of making an error above or below the true value, the technician's 500 results will produce a normal distribution curve whose mean will be at 100 mg. per cent.

Suppose that this same technician were again given 500 specimens of one ml. blood, taken this time from 500 apparently normal persons. Ifthe test were completely accurate, the results of the blood sugar estimations will form a distribution curve around the mean value, which we will again take as 100 mg. per cent.

Let us say the blood sugar test has a standard error of 10 mg. per cent. Then the actual distribution curve of results obtained from the 500 apparently normal persons will be wider than the true distribution curve obtained by a theoretical test with no error, although the two means will coincide. The standard deviation of this wider distribution curve will be approximately the sum of the standard deviation of the true values and the standard deviation of the test. (It has been assumed that the values from the 500 apparently normal persons will produce a fairly normal distribution curve. Even if this happens to be a completely false assumption, it will not alter

the basic conclusion of this section.)

Consider again the previously mentioned population of 10,000 normals and 150 diabetics, being screened by a blood sugar test with a certain standard error. The means of the "normal" and diabetic distribution curves will not change in value, no matter how great the error of the test. Suppose that 130 mg. per cent is used as the screening level. Then the greater the standard error of the blood sugar test, the greater will be the spread of the results for the diabetics, and the greater will be the proportion of diabetics with results below 130 mg. per cent. Similarly the larger standard error of the test will result in a greater spread of the results for the normals, with a higher proportion classified above 130 mg. per cent.

It is therefore concluded that the greater the error of the test, the greater will be the overlap of the normal and abnormal distribution curves, and the greater will be the proportion of false positives and false negatives.

Screening Level and Limits of Normal.

The clinician regards as needing treatment a patient who comes to him with signs and symptoms of anaemia and whose haemoglobin level is below the lower limit of normal. Yet a person who comes for screening, who is below an accepted lower limit of normal, but who is without signs and symptoms of anaemia, may not necessarily have a condition requiring treatment.

Moreover, since the screening test will have a certain

error, a proportion of persons who are actually above the screening level will be wrongly classified as below this level. Since the persons exposed to this risk are the "normal" persons, who are much more numerous than the diseased, a relatively large number will have false positive results, causing a high ratio of false positives to true positives. For these reasons, screening levels cannot be chosen merely by obtaining the "limits of normal" values from a medical text. Indeed there is often no general agreement on those values, which are often derived from small numbers of cases.

The decision to raise or lower the screening level will depend on the seriousness of missing cases of the disease and of referring numbers of false positives. Apart from tuberculosis and carcinoma, the conditions at which screening tests are presently aimed are not sufficiently serious to justify the referral of large numbers of false positives to avoid missing a small proportion of the cases. The screening levels can usually be set on the abnormal side of the "limits of normal."

However, many clinicians will disagree with this method of setting the screening levels. Insurance company statistics show that the values associated with optimum survival are often considerably above the lower limits of normal. If it is the purpose of screening to improve the health of the screenees so that they may lead long and active lives, perhaps the screening levels used should be closer to the "ideal" values associated with optimum survival. This procedure is not a practical

one at present. The referral of all persons above ideal weight or above ideal blood pressure levels will result in referring the majority of apparently well persons taking the tests.

Furthermore, life insurance company statistics cannot be completely accepted at present. They provide only slight evidence, for instance, that when the blood pressure of a hypertensive person is brought down to the "ideal" level this individual's chances of survival will approach the optimum value. In brief, there is little evidence to show that a hypertensive person, under treatment which brings his systolic pressure down to 130 mm. Hg., has the same life expectancy as the normotensive person with the same pressure.

It is now possible to present a number of conclusions from this section. Since they appear to be self evident in their truth, they have been labelled "screening axioms."

Screening Axioms.

When apparently well persons are screened for a condition, they can theoretically be divided into at least two groups:

1. Those who will usually be diagnosed as not having the condition.
2. Those who will usually be diagnosed as having the condition.

When separate distribution curves are drawn for the test results obtained from each group, the curves will overlap to a varying extent.

No matter what screening level is used, the test results must

inevitably include false positives or false negatives or both.

Standardising the conditions under which the screening test is carried out will reduce the overlap of the distribution curves, with a corresponding reduction in false positives and false negatives.

The greater the standard error of the screening test, the greater will be the overlap of the distribution curves, and the greater the proportion of false positives and false negatives.

The so-called "limits of normal" used for diagnostic tests are unsuitable for use as screening levels. The values for screening levels will usually have to be placed on the abnormal side of these limits.

General References

Best,C.H., and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. Published by the Williams and Wilkins Co., Baltimore, Md.

Master,A.M., Garfield,C.I., and Walters,M.B. (1952). Normal Blood Pressure and Hypertension. Published by Lea and Febiger, Philadelphia, Pa.

Wintrobe,M.M. (1951). Clinical Hematology. Published by Lea and Febiger, Philadelphia, Pa.

Wright,S. Applied Physiology. Published by the Oxford University Press, 1952.

SECTION B

THE PLANNING, ADMINISTRATION AND GENERAL RESULTS OF THE BALTIMORE CLINIC

Section B
Chapter 1

THE COMMISSION ON CHRONIC ILLNESS AND ITS BALTIMORE STUDY

Section B
Chapter 1

THE COMMISSION ON CHRONIC ILLNESS AND ITS BALTIMORE STUDY

The Commission on Chronic Illness in the United States was founded in 1949 as a temporary national, non-governmental, research and educational organization. This body consists of a Commission of lay and professional leaders who meet annually to review all the information gathered by the staff, a smaller Executive Committee which is responsible for the administration of the Commission's affairs between the annual meetings, and a full time paid staff which carries out the program adopted by the Commission.

The Commission has listed its goals as follows (Commission on Chronic Illness, undated):

Goal 1. Define the problems arising from chronic illness in all age groups.

Goal 2. Pave the way for dynamic programs to:
prevent chronic illness
minimise its disabling effects
restore its victims to a socially useful and economically productive place in the community.

Goal 3. Clarify the interrelationships of the many professional groups and agencies working in the field.

Goal 4. Coordinate the separate programs for specific diseases with a general program designed to meet more effectively the needs common to all the chronically ill.

Goal 5. Stimulate in every locality a well rounded plan for the prevention and control of chronic disease and for the care and rehabilitation of the chronically ill.

Goal 6. Modify society's attitude that chronic illness is hopeless.

To provide an adequate background for the Commission's recommendations, the chronic illness problem needed more accurate definition. In 1952 the Commission selected Baltimore City as its urban area of study. In the following year the Commission began its Baltimore survey, which consisted of the following five steps (Roberts, 1953):

Step 1. An interview with a responsible member of each of 4,000 households, chosen at random from the total population of the city. Approximately 11,500 persons belonged to these households.

Step 2. Using the household interview data, each person was placed in one of three categories (Commission on Chronic Illness, 1954):

I. Persons reported to have maximum disability.

II. Persons reported to have disease, but less disability.

III. Persons reported to have no disease, or only certain minor conditions.

Each of these categories had been clearly defined in advance, with group II divided into a number of subgroups. A previously determined percentage of persons in each category and subgroup was drawn to produce a subsample of 1,350 persons. The percentages used varied from 100 per cent for the seriously disabled to 6 per cent for those in group III. The private

physicians and hospitals which had taken care of these persons in the recent past furnished additional medical data on these persons.

Step 3. A diagnostic examination and evaluation of this same subsample of 1,350 persons was carried out. While this interview and physical examination was being performed in Johns Hopkins Hospital in Baltimore, each person was given a number of tests which were in many ways comparable to the screening tests in Step 4.

Step 4. From the original sample of 11,500 persons, those who had not been selected to take part in Step 3 and who were 17 years or over, were invited to attend a multiple screening clinic. This group numbered 6,967 persons.

The Commission's primary purpose in Step 4 was to detect additional cases of chronic illness, to add these cases to those found by household interview and physical examination, and to calculate minimum prevalence figures from the results. A discussion on whether or not this purpose was achieved falls outside the scope of this thesis.

The Commission's secondary purpose was to evaluate multiple screening as a technique for encouraging to seek medical care persons who may have certain conditions that benefit from medical treatment.

Step 5. Some of the persons, believed by the examining physician (in Step 3) to have a rehabilitation potential,

were offered the necessary rehabilitative measures. This was carried out only on those persons eligible for the Vocational Rehabilitation Services of the Maryland State Department of Education.

Funds for the everyday work of the Commission and its staff were donated by thirteen national agencies working in the field of medicine, health and welfare. The finances for the Baltimore study have been provided by the United States Public Health Service. The Commonwealth Fund is financing the final report on the study, to be published late in 1956.

REFERENCES

Commission on Chronic Illness (undated). The Commission on Chronic Illness. 615, N. Wolfe St., Baltimore 5, Md., U.S.A.

Commission on Chronic Illness (1954). Method of Selecting Cases for Clinical Evaluation. Mimeographed report.

Roberts, D.W. (1953). Maryland State M.J. 2, 297.

General References.

Roberts, D.W. (1954). Multiple Screening in the Baltimore Chronic Illness Survey. Baltimore Health News 31, 54.

Roberts, D.W., and Wylie, C.M. (1954). Multiple Screening. Maryland State M.J. 3, 557.

PLANNING AND ADMINISTERING THE MULTIPLE SCREENING STEP

Section B
Chapter 2

Invitation Procedure

A part time health educator was employed on July 1 ... up letters of invitation for the nonrespon..., to ... literature pamphlets, and to suggest what general publi... ... used to inform the community of the study.

The problems for the health educator were unusual... ...ticular. 6,567 persons were to be invited to stu... ... scattered throughout the nearly one million citi... ...knows. They consisted of all races and income gr... ... had no opportunity to volunteer or decline parti... ...in the study.

PLANNING AND ADMINISTERING THE MULTIPLE SCREENING STEP

In March 1954 a preliminary description of the Multiple Screening Step (Commission on Chronic Illness,1954a) was distributed to thirty-five persons in the medical and public health fields. The resulting suggestions and criticisms were incorporated, when believed useful, in the final arrangements for the screening clinic.

In May 1954 a Technical Advisory Committee to the screening step was formed to represent the practicing physicians in Baltimore. Its nine members had two meetings, and made a number of helpful suggestions to prevent possible difficulties with family doctors while the clinic was in progress.

Invitation Procedure

A part time health educator was employed on July 1954 to draw up letters of invitation for the screenees, to design explanatory pamphlets, and to suggest what general publicity should be used to inform the community of the study.

The problems for the health educator were unusual and difficult ones. 6,967 persons were to be invited to attend the clinic, scattered throughout the nearly one million citizens of Baltimore. They consisted of all races and income groups, and had had no opportunity to volunteer or decline participation in the study.

A number of these persons belonged to families where one member had already attended the physical examination clinic,

and had presumably been pleased by their experience. It was therefore decided that an initial personal letter would be written to the "evaluees" who had attended for physical examination, asking their help in persuading the restcof the family to attend for screening. In those families without an evaluee, the initial explanatory letter was addressed to the head of the household. This first letter, individually typed and signed by hand, was mailed four weeks before the appointment date for the members of the family to come for screening. It was accompanied by a pamphlet (Appendix A) which explained the purpose of multiple screening. Special letters were written to suit situations where evaluees were too young to have a persuasive influence or had refused to accept the physical examination.

Two weeks later, the letter of invitation was mailed to the head of the household. This personal letter was again accompanied by a pamphlet (Appendix B) designed to overcome any last-minute fears concerning the procedure. Appointment cards (Appendix C) were enclosed for each member of the family. Each individual was able to change the appointment time if desired, and stamped-addressed postcards were included for this purpose.

General Publicity.

It was decided not to have a vigorous general publicity campaign, which might attract large numbers of persons who were not part of the sample under study. Just before the clinic began, articles were submitted to and published by two of the main Baltimore newspapers and by a number of newspapers serving

local sections of the city.

The October 1954 issue of the Baltimore Health News, widely distributed to physicians and lay persons interested in the health problems of the city, was mainly devoted to a description of the screening program to begin on that month (Roberts, 1954). The Maryland State Medical Journal, in its October issue which went to the majority of practicing physicians in Baltimore, included a more technical article on the multiple screening step (Roberts and Wylie, 1954). Reprints of this article were sent to each physician when he received his first reports from the screening clinic.

Manual of Procedures

A comprehensive manual of procedures was developed to describe all the steps which had to be taken, from the time the first letter of invitation was mailed to obtaining the final follow-up diagnosis from the family doctor. This manual included copies of all letters to be used, a detailed description of the necessarily complicated filing system used to keep track of the individual's progress through the screening procedure, and instructions on how to carry out the individual screening tests (Commission on Chronic Illness, 1954b).

The Clinic Location

In August 1954, it was decided to accept the offer made by the Maryland State Health Department of its auditorium named Bennett Hall. This air-conditioned hall, with an

adjacent corridor which could also be used as clinic space, made a total available area of 2,300 square feet. This location had some parking facilities, and had an elevator available for persons who were unable to climb the two flights of stairs. Adequate lavatory facilities were located just outside the hall.

The Personnel

Two public health nurses, who were to do much of the clinic administrative work, and a clerical supervisor took part in the immediate preparations. Eighteen typists, chosen because of pleasant personality and appearance, started a training period nine days before a trial run of the clinic was scheduled. This training period had been planned to orientate the workers in the purposes of the clinic, and to teach them how to do some of the tests.

The more difficult tests were taught to the two nurses, and to four typists whose abilities appeared well above average. At the end of this brief orientation period, all the workers knew a little about each test. Several knew a lot about one specific test which would be their initial assignment when the clinic began. All were able to give intelligent answers to the screenees' questions about the clinic. Perhaps best of all, the clinic workers had developed a liking for each other and a sympathy with the purpose of the study.

The Clinic Arrangement.

Curtained exhibition booths were installed two days before the initial trial run. The location of these booths and the

general arrangement of the clinic are shown on the plan overleaf. The hearing test was carried out in a soundproof booth installed in the auditorium.

The tests were carried out in the following order:

Height and weight) while taking a glucose drink.
Questionary)

Dental examination. This required the continuous presence of a dentist, supplied mainly by the United States Public Health Service. It was included because of the Commission's interest in the prevalence of dental conditions. It was not strictly regarded as one of the screening tests.

Six-lead electrocardiogram) both taken in the same booth while
Blood pressure) the person was lying down.

Visual acuity test, using the American Optical Company Sight Screener.

Audiometric sweep-check hearing test.

70 mm. chest x-ray.

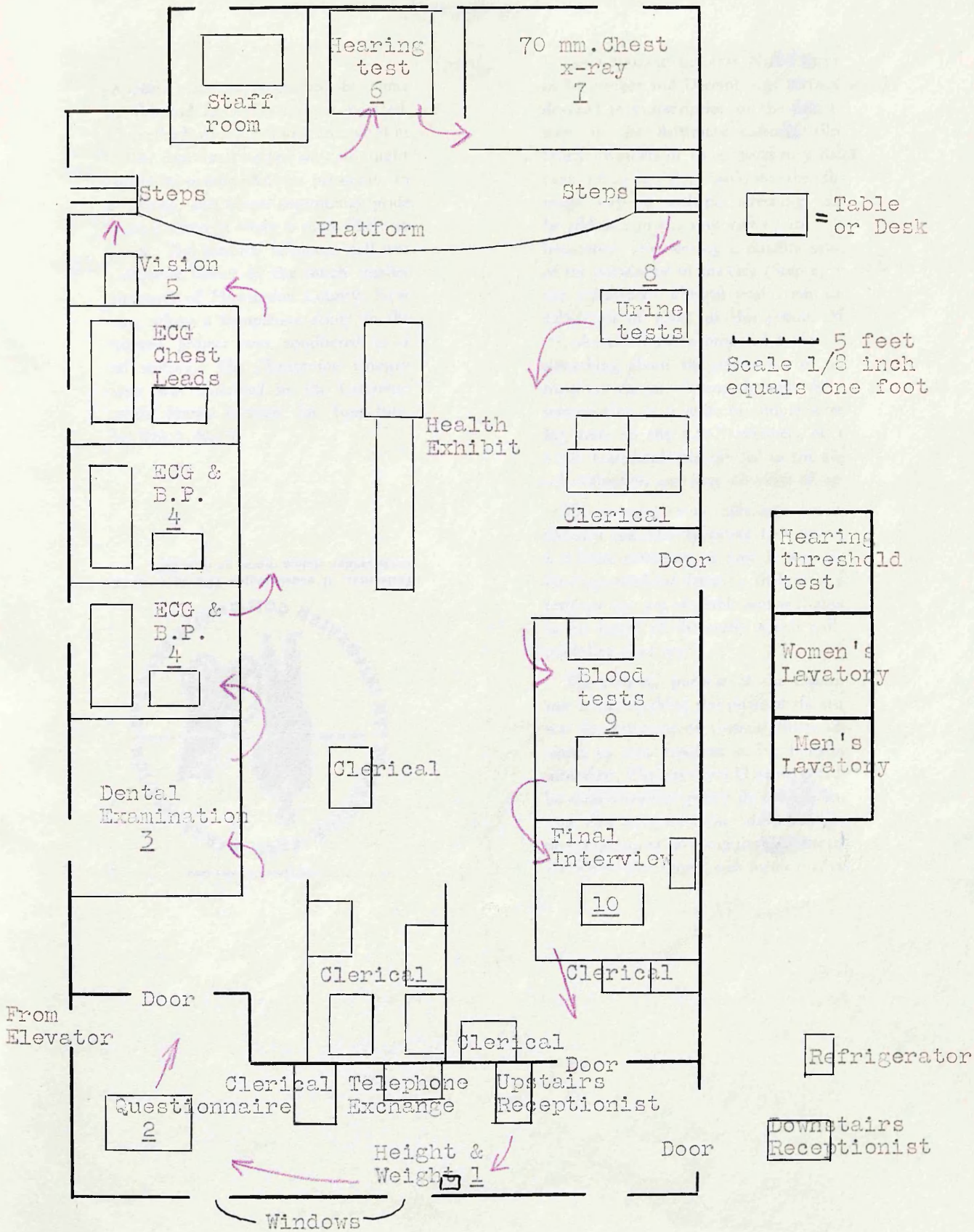
Clinitest and Bumintest for urine sugar and albumin respectively.

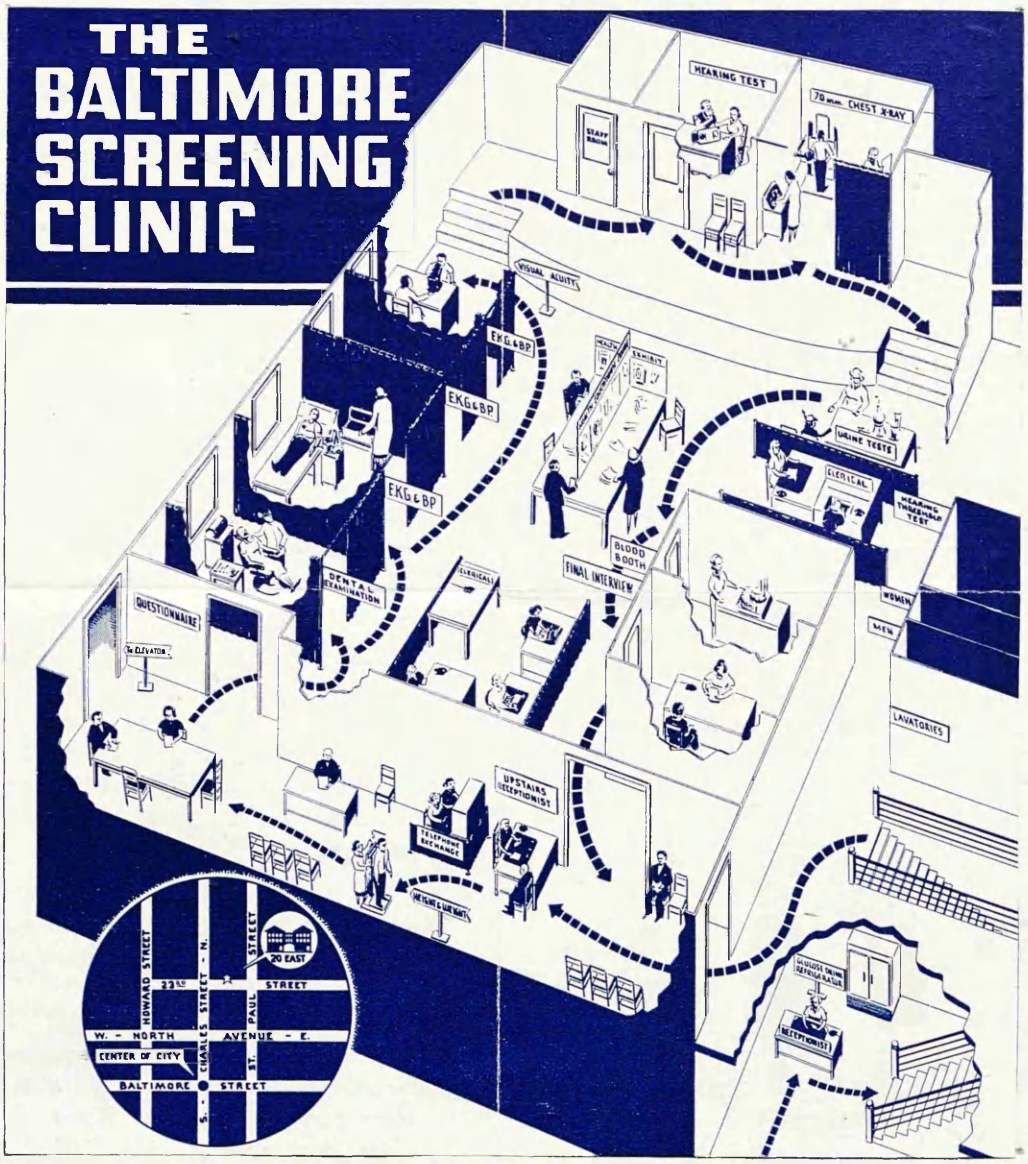
Copper sulphate specific gravity test for "anaemia".

Serological test for syphilis.

Wilkerson-Heftmann test for true blood glucose content, using the Clinitron.

PLAN OF BALTIMORE SCREENING CENTRE




THE
BALTIMORE
SCREENING
CLINIC
HEARING TEST
70 mm. CHEST X-RAY
STAFF ROOM
VISUAL ACUITY
EKG & BP
EKG & BP
EKG & BP
EXHIBIT
URINE TESTS
CLERICAL
HEARING THRESHOLD TEST
BLOOD BOOTH
FINAL INTERVIEW
CLERICAL
DENTAL EXAMINATION
QUESTIONNAIRE
ELEVATOR
WOMEN
MEN
LAVATORIES
UPSTAIRS RECEPTIONIST
TELEPHONE EXCHANGE
WEIGHT & HEIGHT
ILLNESS DURING EXAMINATION
RECEPTIONIST
20 EAST
22ND STREET
HOWARD STREET
CHARLES STREET - N.
PAUL ST.
W. - NORTH AVENUE - E.
CENTER OF CITY
BALTIMORE STREET
S.

REFERENCES

Commission on Chronic Illness. (1954a). Multiple Screening Step (a preliminary description for discussion, criticism and refinement). Mimeographed prospectus.

Commission on Chronic Illness (1954b). Manual of Procedures for Screening Clinic. Mimeographed.

Roberts, D.W. Baltimore Health News 31, 54. (1954).

Roberts, D.W., and Wylie, C.M. Maryland State M.J. 3, 557. (1954).

PROBLEMS OF CLINIC ADMINISTRATION

Section B
Chapter 3

The Trial Run

PROBLEMS OF CLINIC ADMINISTRATION

The Appointment Procedure

It was believed that the clinic could take care of one person every five minutes, since that was the time within which each test could be carried out. Two booths had been installed on the first day for taking the electrocardiograms and blood pressures, since this double procedure was likely to take almost twice as long as the other tests.

It was not feasible to have a clinic, running at full schedule and with screenees in every testing booth, close down completely for lunch. It was therefore arranged that appointments during the lunch hour would be scheduled every ten minutes, with half of the clinic workers eating during the first half hour and the remainder during the second half. Those not eating would carry out two tests instead of their usual one.

In the morning, each employee began work at a time shortly before the first screenee was expected to arrive at her testing station. She had a corresponding time, eight hours later, for finishing in the afternoon.

The Trial Run

Sixty-four city health department workers acted as "screenees" during a trial run held on October 1, 1954. A number of technical defects were corrected as a result of this valuable experience. The two persistent bottlenecks were corrected by installing a third ECG booth, and by rearranging the blood booth so that two "blood takers" could comfortably work at

the same time. In the second and third week of the clinic, two persons with experience in taking blood became available, ending the personnel difficulties in the blood station. The difficulty with the chest x-rays, which were quite unsatisfactory during the trial run, was overcome when a trained radiographer was employed several weeks after the clinic began.

Choice of "Personal Physician"

Each screenee was asked to write down on the questionary the name and address of the physician to whom the results should be sent. About 20 per cent of the screenees did not have a definite family physician. These persons were asked to choose one from a list of Baltimore physicians taken from the telephone directory. A small number did not make this choice until they had returned home and discussed the problem with their friends and family. A significant number of the positive screenees were followed-up by physicians whom they had not previously visited.

Courteous Treatment of the Screenee.

During the screenee's visit to the clinic, he was treated with utmost courtesy and introduced to each technician on arrival at the new testing station. This produced a friendly and relaxed atmosphere which received much favourable comment from the screenees. It may have been the deciding factor in many instances when one member of the family visited the clinic to "spy out the land", and returned home to persuade the rest of the members to come.

The screening test results were recorded in code form

on the appointment card which the screenee carried with him. Once the blood sugar had been taken, he waited an additional ten minutes for the blood sugar result, and was then interviewed by one of the two public health nurses before leaving. All the results, apart from the ECG reading, chest x-ray and STS, were thus available to the nurse during the interview.

The Nursing Interview

The primary purpose of the interview was to prepare the screenee, in so far as was possible, for the type of follow-up letter he was likely to receive three or four weeks later. The secondary purposes were to satisfy any questions the screenee might have about the procedure and results, to see if he could help persuade any other invited members of his family to attend, and to try to ensure that no grossly evident conditions were present which could not be detected by the screening procedure.

The interviews, taking from five to ten minutes, appeared to be very satisfying to the screenees. Almost all of them left feeling that their questions had been satisfied, and that if their letter advised them to visit their personal physician, they would be well advised to do so. As was expected, many screenees with emotional problems took advantage of the interview to bring their worries out into the open.

In brief, the interviews seemed helpful, not only in improving public relations, but also in reducing the need for home visits to persuade positive screenees to visit their family physician.

Attendance Problems

A mailing firm used electric typewriters to write the

body of the personal invitation letter, and the name and address were typed by hand. This personal approach had been planned in the belief that mimeographed letters would be largely ignored by the persons receiving them.

Many difficulties prevented the scheme from working efficiently. There was some delay in sending the names and addresses to the mailing firm. This firm found that the large amount of work involved prevented them from keeping up to the arranged schedule.

Attendance was very low during the first two weeks. About the third week, it was discovered that many individuals were receiving the invitation letters just before the invitation date, and in some cases after that date had passed.

During the next week, while attempts were being made to speed up the process, it was decided to allow volunteers from the general population to fill in gaps in the clinic schedule. Even with this alteration, however, the clinic was operating at about one-third of its full capacity for long periods of time. Since the expense of the clinic was mainly involved in salaries, this greatly increased the apparent cost of screening each person.

The pamphlets which accompanied the personal letters, generally agreed to be attractive and simple by the health educators who criticised them, were not found to be universally helpful. Indeed, a number of screenees coming in at a later date commented that whenever they saw a pamphlet, they tore up the letter believing it to be an appeal for funds.

All persons who failed to respond to the original invitation received a mimeographed letter of invitation at a later date. This letter was sent by a second mailing firm which also, at one period of several days, got so far behind with the procedure that the letters were arriving after the appointment date.

Those persons who failed to respond to the mimeographed letter were contacted by phone if they had a phone. The calls were made by typists with pleasant telephone personalities, who were briefed on all the awkward questions which were likely to be asked. A significant improvement in response to the last letter of invitation occurred when the screenees were phoned to remind them of their appointment date. Those individuals who did not have a phone were visited at home by five home interviewers, four of whom had taken part in the initial home interviewing step in this study.

Arrangements were made with a local taxi company to pay for transporting persons who regarded this problem as being the main barrier to their attendance. This group turned out to be very small in number.

General publicity was increased about the fourth week, and included a somewhat sensationalised newspaper article. This article resulted in several hundred phone calls and letters from persons, not in the sample being studied, who wished to be substituted for those who did not appear. Screenees attending at a later date specifically mentioned this article, however, as the first time they had noticed any mention of the clinic,

and as the article which aroused their interest in the procedure.

About the fifth to eighth weeks of the clinic, one radio and four television programs, all aimed at the housewives of Baltimore, devoted part of their programs to discussing the clinic. While many screenees commented that they had seen one of these programs, it is not known whether they helped deliver a significant number of persons.

Mimeographed Versus Individually Typed Letters.

During two days in the eighth week of the clinic, 61 families, involving 117 individuals, were invited in by the usual typed letter. At the same time, 30 families, involving 61 individuals, were invited in by an identical letter, except that it was mimeographed instead of individually typed. (This experiment was devised and supervised by Dr. Herbert Caron of the United States Public Health Service, and by Mr Dean Krueger of the Commission on Chronic Illness, by whose kind permission these results are quoted).

The two groups of invitees were chosen at random, and were believed to be fairly similar in all respects likely to affect their response to invitation. Members of both groups who did not respond to the first invitation received the routine secondary persuasive effort -- second mimeographed letter of invitation, phone call or home visit.

Letter Used	Number of Families	Percent of Families Attending	Number of Individuals	Percent of Individuals Attending
Typed	61	36%	117	27%
Mimeographed	30	47%	61	39%

The attendance of the two groups is shown in the preceeding table. A family is regarded as having responded when any one member attended for screening. 27 per cent of the individuals receiving the typed letters ultimately attended for screening. (This is closely similar to the attendance of the total 6,967 invitees, of whom 29 per cent responded). Contrary to expectation, there was a greater response of 39 per cent of those persons receiving the mimeographed letter.

This difference does not justify the statement that the mimeographed letter was significantly more effective than the typed letter. It does appear reasonable to say, however, that the mimeographed letter was not markedly less effective than the typed letter. Since $1,556.00 was spent in this study on the invitation letters, and since the typed letter was twice as expensive as the mimeographed letter, this matter should certainly receive further study.

In discussing this result, Caron (1955) suggests:

"Since a mimeographed latter is probably interpreted by most respondents as less direct and personal than a typed latter, the recipient of a mimeographed letter may be less likely to feel himself as 'singled out' for direct contact. The individual who perceives himself as 'singled out' may be more prone to perceive the letter in terms of 'promotional literature'...

"In addition, perhaps the typed letter (more than the mimeographed) suggests that money has been invested and will have to be regained in some form.

"A mimeographed letter may also suggest a lengthy and

anonymous list of recipients (more than with a typed letter). Individuals desiring a health examination, but also desiring anonymity for any reason, may feel more reassured in responding to the mimeographed letter..."

The Clinic Morale

The nine day period of indoctrination of the clinic workers resulted in a remarkable amount of enthusiasm for the job. The most morale-straining periods occurred when the invitation procedure was in difficulties and attendance was poor. The workers, with much faith in the value of the clinic, were quite disappointed at the apparent lack of public interest.

Because all of them had typing experience, however, it was possible to keep most of them fairly well occupied during the slack periods. The weekly staff conference, attended by all the clerical and technical workers, appeared to help greatly in preventing the sagging of the morale. At these meetings, which sometimes were only ten minutes long when screenees were still in the process of going through, the latest problems of the clinic were discussed. Many of the workers had useful suggestions on how the procedures could be improved, and these improvements were often carried out on the following day. The meetings were quite informal, were participated in by all members of the staff, and appeared to result in quite a marked uplift in the morale on the following day.

Personnel Requirements.

The maximum number of personnel used at any one time was as follows:

9 clerk-typists; 10 untrained technicians; 3 trained technicians; 2 "phone-callers"; 5 home interviewers; 2 public health nurses; 1 clerical supervisor; 1 physician; 1 part time janitor.

Physician Time Involved

A physician was constantly present during the clinic sessions. If the program had been continuous, however, it is likely that a public health nurse would have taken over much of his work. Professional time was also involved in interpreting the ECG and chest x-ray. No time studies were carried out, but it is known that these readings occupied the physicians for only a small proportion of each day.

REFERENCE

Caron, Herbert (1955). Participation in Screening Surveys. Report of an Experiment. Unpublished.

THE CHARACTERISTICS OF THE SCREENEES

Section B
Chapter 4

THE CHARACTERISTICS OF THE SCREENEES

The Commission invited 6,967 persons to attend for screening. This group consisted of those persons in the original sample of 11,500 who had not been selected to take part in the evaluation procedure (Step 3) and who were 17 years and over. The minimum age of 17 years was selected for two reasons:

1. The prevalence of the conditions at which the tests were aimed was likely to be quite low in those under 17 years.
2. Multiple screening appeared less necessary and less suitable for those under 17 years, since well-baby and preschool clinics and school health programs usually care for their health needs in the United States.

Table 1 - Screenees Classified by Age, Race and Sex.

Age in Years	TOTAL 1/			WHITE			NON-WHITE		
	Both Sexes	Male	Female	Both Sexes	Male	Female	Both Sexes	Male	Female
All ages	2024	922	1086	1562	734	828	429	177	252
	Percent distribution 2/								
All ages	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
17-24	11.0	9.5	12.2	9.4	8.4	10.3	17.0	14.7	18.7
25-34	25.0	23.2	26.6	24.1	21.4	26.4	28.4	30.5	27.0
35-44	27.5	29.8	25.4	28.6	31.6	25.8	24.0	23.7	24.2
45-54	18.5	19.0	18.1	19.0	19.3	18.6	16.8	16.4	17.1
55-64	11.5	12.3	10.9	12.2	13.2	11.4	8.9	8.5	9.1
65-74	5.4	5.0	5.8	5.7	4.9	6.4	4.0	4.5	3.6
75or over	1.1	1.2	1.0	1.1	1.1	1.1	0.9	1.7	0.4

1/ Includes 33 persons with age, sex or race unknown
2/ Excludes those of unknown age.

This chapter will describe the age, sex and racial characteristics of the screenees. Their characteristics will be compared with those of the Baltimore population and with those of the total group invited to attend for screening. Table 1 shows the age, sex and racial classification of the persons screened.

Age. Fifty-two per cent of the screenees were between the ages of 25-44 years.

Sex. Fifty-four per cent of the screenees were women. The women were slightly younger than the men, 39 per cent being under 35 years of age compared with 33 per cent of the men. White screenees had a lower proportion of women -- 53 per cent as compared with 59 per cent among the non-white screenees.

Race. The non-white screenees were younger than the white screenees. Forty-five per cent of the non-white screenees were under 35 years compared to 34 per cent of the white screenees.

Table 2 of this chapter compares the age and racial characteristics of the screenees with those of the Baltimore population. In general, the screenees were younger than the Baltimore population, largely due to the greater attendance of white screenees of 44 years and younger. There was relatively little difference in the age distributions of the non-white screenees and the non-white population of Baltimore.

Table 3 of this chapter shows the percentage of those invited who attended the clinic. In general, the proportion of white invitees attending reached a maximum in the 35-44 year age group, falling steadily in the younger and older age groups.

Table 2 - Comparison of Age and Racial Characteristics of Screenees and Baltimore Population (Percent Distribution)

Age in Years	BALTIMORE POPULATION 1/				SCREENEES 2/			
	Total	White	Non-White	Percent non-white	Total	White	Non-White	Percent non-white
All ages	100.0	100.0	100.0	24.4	100.0	100.0	100.0	21.6
17-24	13.3	12.6	15.4	28.2	11.0	9.4	17.0	33.2
25-34	23.4	22.3	27.0	28.1	25.0	24.1	28.4	24.4
35-44	21.7	20.8	24.4	27.5	27.5	28.6	24.0	18.7
45-54	18.0	18.1	17.7	24.0	18.5	19.0	16.8	19.4
55-64	12.9	14.1	9.2	17.4	11.5	12.2	8.9	16.5
65-74	7.5	8.4	4.6	14.9	5.4	5.7	4.0	15.6
75 or over	3.3	3.8	1.8	13.7	1.1	1.1	0.9	18.2

1/ Population July 1, 1954, as estimated by the Baltimore City Health Department. 2/ Excludes unknown race.

Table 3 - Percentage of Invitees who attended the Screening Clinic, by Age, Sex and Race.

Age in Years	TOTAL			WHITE			NON-WHITE		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
All ages	29.1	28.7	29.3	30.2	30.7	29.8	24.5	21.6	26.9
17-24	21.7	22.0	21.6	20.4	21.8	19.5	25.1	22.4	26.9
25-34	30.8	28.2	33.1	33.2	29.2	36.9	24.3	24.2	24.4
35-44	35.0	35.4	34.5	38.7	40.3	37.1	24.2	20.8	27.4
45-54	29.9	28.0	31.8	30.7	29.6	31.9	25.4	19.9	31.4
55-64	27.6	29.0	26.3	28.1	31.7	25.2	23.9	18.1	30.3
65-74	23.5	23.5	23.6	23.2	22.6	23.7	21.2	21.6	20.9
75 or over	12.2	17.2	9.4	10.8	14.8	8.7	16.7	30.0	7.1

There were no consistent differences in the attendance of white men and women. Non-white women had a better attendance than non-white men. The improved attendance of the white group was much more marked in the men than in the women.

Summary

The invitation procedure used in the Baltimore clinic produced a better response in white than in non-white persons, and in those aged 25 to 54 years inclusively than in those below and above this age range.

SUMMARY OF RESULTS OF THE BALTIMORE CLINIC

Section B
Chapter 5

SUMMARY OF RESULTS OF THE BALTIMORE CLINIC

This chapter will briefly present the results of the screening tests, to be followed by a detailed presentation and discussion of the results in Section C of this thesis.

Table 1 - Screenees Classified by the Number of "Abnormalities" Found.

	Number	Percent of Total screened	Percent of Total with major abnormality
Total persons screened	2024	100.0	
No abnormalities	743	36.7	
"Minor" abnormalities only	629	31.1	
One or more "major" abnormalities	652	32.2	100.0
One major abnormality	434	21.4	66.6
Two major abnormalities	140	6.9	21.5
Three major abnormalities	55	22.7	8.4
Four or more major abnormalities	23	1.1	3.5

Thirty-seven per cent of all screenees had no abnormal test results, as shown in Table 1. An additional 31 per cent were regarded as having 'minor abnormalities' only. These persons had positive results only in the tests for obesity, hearing, vision or glycosuria (without hyperglycaemia). Thirty-two per cent of the screenees failed in one or more of the remaining tests. Of these latter 652 persons, 45 had answered in the questionary that they were already aware of the condition at which the positive test was aimed. The remaining

607 individuals were asked to visit their family physicians for further advice.

To obtain the follow-up results, the staff got in contact with the 607 screenees by mail, or by phone or home visit when necessary, to find out if they had visited their physicians. Sixty-five per cent (393) reported that they had consulted their physicians about the results. In turn, the physicians of these 393 individuals were contacted. The necessary information was obtained, by mail or by phone, from the physicians of 351 of the referrals.

Tables 2 and 3 of this chapter summarise the results of the screening tests and of the follow-up procedure. The number of persons taking each test is not constant (column 1), since one of the tests was sometimes not done on an individual or the result was unsatisfactory for various reasons. For instance, a considerable proportion of unsatisfactory chest x-rays in the early weeks of the clinic resulted in a smaller number of screenees completing this test. A number failed to answer both questions regarding symptoms of "heart disease". A small number of persons did not wish to take the urine or blood tests. The remaining tests were completed in a high proportion of screenees.

When the screenee marked on the screening questionary that he knew he had the condition at which the test was aimed, the positive result was classified as previously known. When no mark was made concerning the condition, the result was classified as previously unknown (column 3) and the screenee

Table 2 - Summary of the Screening Test Results and of the Follow-up Procedure

Condition for which screened	Persons screened (1)	POSITIVE RESULTS Total (2)	POSITIVE RESULTS Previously unknown to screenee (3)	Follow-up completed with physician (4)	CONDITION CONFIRMED Total (5)	CONDITION CONFIRMED Previously unknown to physician (6)
Hypertension	2021	150	67	41	31	11
Heart Dis. (ECG)	2020	247	194	128	43	16
Heart Dis. (X-ray)	1767	182	155	67	29	8
Heart Dis. (Questionary)	1898	113	78	42	17	5
Heart Dis. (3 tests)	2020	428	357	173	67	23
Tuberculosis	1767	34	28	16	2	2
Other Chest	1767	64	50	17	2	1
Proteinuria	1946	88	79	37	2	2
Diabetes	1916	15	14	9	5	4
Anaemia	1980	32	29	13	8	7
Syphilis	1949	53	*	14	4	1
Obesity	2021	106 abnormal, and previously unaware of condition				
Impaired Vision	2006	530 abnormal				
Impaired Hearing	2016	76 abnormal, and previously unaware of condition				

*Screenees were not asked if they already knew they had syphilis.
Note: Column (1) excludes persons not receiving a test and those for whom a test result was unsatisfactory.

Table 3 - Summary of the Screening Test Results and of the Follow-up Procedure, expressed in Rates per 1,000 persons screened.

Condition for which screened	Rate per 1,000 persons screened					
		POSITIVE RESULTS			CONDITION CONFIRMED	
	Persons screened (1)	Total (2)	Previously unknown to screenee (3)	Follow-up completed with physician (4)	Total (5)	Previously unknown to physician (6)
Hypertension	1000	74	33	20	15	5
Heart Dis. (ECG)	1000	122	96	63	21	8
Heart Dis. (X-ray)	1000	103	88	38	16	5
Heart Dis. (Questionary)	1000	60	41	22	9	3
Heart Dis. (3 tests)	1000	212	177	86	33	11
Tuberculosis	1000	19	16	9	1	1
Other Chest	1000	36	28	10	1	1
Proteinuria	1000	45	41	19	1	1
Diabetes	1000	8	7	5	3	2
Anaemia	1000	16	15	7	4	4
Syphilis	1000	27	*	7	2	1
Obesity	1000	52 abnormal, and previously unaware of condition				
Impaired Vision	1000	264 abnormal				
Impaired Hearing	1000	37 abnormal, and previously unaware of condition				

* Screenees were not asked if they already knew they had syphilis.
Note: Column (1) excludes persons not receiving a test and those for whom a test result was unsatisfactory.

referred for diagnostic examination. However, the physicians subsequently stated that a considerable proportion of the confirmed cases were previously known to them. All conditions classified (in column 6) as previously unknown were clearly and definitely stated by the physicians to be so.

The marked conflict between the statements of the screenees and the physicians on whether the conditions were previously known has three possible explanations:

1. The screenee may not have been told by the physician that the condition was present, or may have misunderstood what the physician had said.
2. The screenee may have known of the condition, but have failed to mark it on the screening questionary.
3. The physician may have misinterpreted the questionary mailed to him, and have mistakenly marked the condition as previously known when, in fact, it was previously unknown.

Whatever the true explanation, however, the physicians' replies were used to classify the confirmed conditions and to determine the "yield" of each test in terms of previously unknown, confirmed cases (column 6).

SECTION C

DISCUSSION AND DETAILED RESULTS OF THE BALTIMORE CLINIC

SCREENING FOR DIABETES MELLITUS

Section C
Chapter 1

SCREENING FOR DIABETES MELLITUS

At autopsy cases of diabetes mellitus frequently, but not invariably show hydropic degeneration in the beta cells of the islets of Langerhans. Animal experiments have shown (Copp and Barclay, 1923; Lukens and Dohan, 1940; Lukens et al, 1943) that control of the diabetes prevents this progressive change in the cells, and that cells which already show this change can be rescued by controlling diabetes with insulin and diet. Lukens et al.(1943) believe that the hyperglycaemia is responsible for the degenerative changes in the islets.

Allen (1952) holds that the results of such studies are comparable with the progressive worsening of the human case of diabetes which frequently occurs when the disease is uncontrolled. He states that diabetes progresses as long as the blood sugar remains high, even without glycosuria.

At present the treatment of diabetes begins too late to expect restoration of the pancreatic islet functions, states Wilder (1948). Both he and Wilkerson et al. (1949, 1955) believe that treatment is most effective when begun at the very onset of the disease.

The cardiovascular-renal complications of diabetes show a definite correlation with the duration of the disease (Parkhurst and Betsch, 1955), but are less well correlated with its lack of control (Rickets, 1954). Dolger (1947) comments that these changes may be an essential element in the disease and independent of its treatment. Collens (1954) suggests

that the complications may occur because an insulin deficiency is only a part of the diabetic syndrome, in which other enzymatic and endocrine factors play a role. However, Dunlop's (1954) studies in Edinburgh suggest -- although the evidence is not conclusive -- that aggressively treating the disorder is most important in preventing or postponing these complications.

Degenerative changes are occasionally present when diabetes is first discovered (Ditzel and Sagild, 1954). This finding suggests that a long presymptomatic period of hyperglycaemia may be the cause of the complications under strict control (Parkhurst and Betsch, 1955). Until follow-up studies are made on diabetics found in the early stage by detection clinics, this possibility cannot be confirmed or denied.

Conflicting Philosophies of Treatment.

As a result of the failure to prove conclusively that strict control of diabetes is essential, the "free diet" method of treatment has gained some popularity. Tolstoi (1953; Tolstoi et al, 1939, 1950) popularised this method in the United States in 1940, and enumerated his criteria for adequate control as follows:

1. Eliminate all symptoms of diabetes.
2. Maintain or increase weight.
3. Strictly avoid ketonuria.
4. Disregard glycosuria.
5. Restore patient to social and economic usefulness.

Diabetic patients in the presymptomatic early stage already fulfill the therapeutic aims of this method. These persons are therefore unlikely to benefit by their early referral to physicians of the "free diet" school until the disease advances further.

Probably the majority of physicians in the United States adhere more closely to Joslin's (1952) "regulated diet" method of treatment, which aims at keeping the urine sugar-free for much, if not all of the time. This school recommends the early detection and vigorous treatment of diabetes, using carefully regulated diet and insulin dosage.

The Prediabetic State

The potential diabetic has been described by Joslin (1952) as a person whose glycosuria is closely related to the diet, and who easily becomes sugar-free with slight dietary restrictions. The blood sugar values are slightly below the arbitrary values set for diagnosing diabetes.

At present there is no evidence that the potential diabetic will inevitably develop diabetes or even that he is likely to do so. Indeed, there is no general agreement that a prediabetic state exists. However, Jackson (1955) goes so far as to suggest that in the prediabetic state the disorder which will eventually lead to hyperglycaemia is already making itself known. He feels that the conditions should also be called "diabetes", on the basis that the disease is already there, with its vascular changes and other possible complications.

Glycosuria without Hyperglycaemia.

Joslin (1952) has stated that ten per cent of cases of low renal threshold glycosuria eventually become true diabetics. Other clinicians frequently quote this statement, although there are no extensive follow-up studies to support it.

Low renal threshold glycosuria (Joslin, 1952) needs no treatment. Until there is more definite evidence that such persons are potential diabetics, there is no valid reason why diabetes detection clinics should refer them to their family physicians.

Glycosuria in Pregnancy.

Glycosuria is fairly common in pregnancy (Cantarow and Trumper, 1955), the frequency varying from 5 to 40 per cent. This glycosuria is not usually associated with hyperglycaemia (Bodansky, 1952), and is apparently not an indication for treatment. However, screening clinics should refer to their personal physicians all pregnant women with hyperglycaemia as well as glycosuria.

Hyperglycaemia in the Elderly

A raised blood sugar without glycosuria, due to a high renal threshold, is more common in older persons and does not warrant the diagnosis of diabetes (Joslin, 1952; Wagner, 1955). Glucose tolerance tests of elderly subjects (Deren, 1937; Chesrow and Bleyer, 1954) show a delayed peak and prolongation of the curve, frequently accompanied by a high renal threshold. Higher screening levels for persons over 60 years seem indicated in diabetes detection programs to avoid referring too

many false positives. This procedure was not carried out, however, in the Baltimore clinic.

Other Factors Affecting Blood Sugar Values

Loughlin et al.(1943) state that the tourniquet used in collecting venous blood produces immediate and significant differences in the venous blood sugar, amounting to 20 or 25 mg. per cent. The evidence for this finding, however, is not convincing.

Mosenthal (1947) states that menstruation is accompanied by lowered glucose tolerance. There are not sufficient data, however, to suggest that diabetes screening levels should be changed for menstruating women.

Diabetes and Obesity

"Obesity has a deleterious effect on carbohydrate metabolism and appears to impair glucose tolerance even in non-diabetics. The longer the duration of the obesity, the greater is this impairment." (Wolfson, 1954). A number of studies confirm this association (Mattison, 1955). Contradictory opinions are held, however, on the success of weight reduction in improving the status of obese diabetics (Newburgh and Conn,1939; Richardson, 1953).

The familial tendency to diabetes is well established (Pincus and White, 1952). The possibility that obesity may make manifest a latent diabetic tendency would suggest that weight control is especially important in persons with close diabetic relatives. Wilkerson et al. (1955) recommend that any community-wide diabetes detection program should include special appeals to such persons.

Blood Sugar Preservation

Enzymes present in blood cause the glucose to disappear at an approximate rate of five per cent per hour at room temperature (Sunderman, 1951). The most promising agent for inhibiting the action of these enzymes is a combination of 1 mg. of thymol with 10 mg. of sodium fluoride per millilitre of whole blood (Sander, 1923). This preservative permits specimens to be stored or shipped for periods up to 96 hours without showing any significant change in glucose content (Bowman and Enterline, 1954). It is thus quite feasible for diabetes detection clinics to be held some distance away from the point at which the blood glucose estimations are being carried out.

Capillary versus Venous Blood

The use of capillary blood in diabetes screening programs has several advantages. Persons capable of doing venipunctures are not needed. No container is needed for the blood, which is taken directly into the measuring pipette. The practical disadvantages are that there is no reserve of blood to repeat the test if the first estimation is not successful, or to carry out additional tests for anaemia and syphilis.

Wilkerson et al. (1955) favour the use of venous blood in diabetes screening, "since it is less variable, and the results obtained are easily interpreted and more readily accepted" in the United States. Even more important is the fact that venous blood glucose values taken after ingestion of food or glucose show much greater differences between diabetic persons

and non diabetics than do capillary blood values (Cantarow and Trumper, 1955). The use of venous blood in screening for diabetes will therefore improve the specificity of the test.

Advantage of True Glucose Values

The Folin and Wu method, or its modifications, is the most commonly used blood sugar test. This procedure measures the total reducing substances in the blood -- glucose plus variable amounts of glutathione, fructose, cysteine, ergothioneine, creatinine and other undetermined materials. Although these materials were believed to give a fairly constant figure of 20-30 mg. per cent (Todd et al.,1953), there is evidence that they may vary from one mg. per cent to over 70 mg. per cent (Haunz and Keranen, 1950; Mosenthal and Barry, 1946), and that they may not be constant even in the same person from hour to hour.

Large amounts of these substances may therefore cause non diabetic persons to have Folin and Wu values in the diabetic range. The specificity of the blood glucose screening test will be improved by using a procedure which does not measure non glucose reducing substances. One of these "true glucose" tests is the Wilkerson-Heftmann test (Wilkerson and Heftmann, 1948), which was used in the Baltimore clinic.

Problems in Diagnosing Diabetes

Leading clinicians in the United States do not agree on several points in the diagnosis of diabetes. Joslin (1952) believes that glycosuria must be present, but not necessarily persistent, for the diagnosis to be made. This recommendation

was followed in the Baltimore clinic, where all screenees referred as "probable diabetics" were required to have glycosuria in addition to positive blood sugar tests. It was reasoned that even intermittent glycosuria was likely to show up following the ingestion of 50 gm. glucose on arrival at the clinic.

In retrospect, this decision may have caused several diabetics with high renal thresholds to be missed because of negative urines. Duncan (1951) and Wilkerson et al. (1955) believe that glycosuria need not be present for a diagnosis of diabetes to be made.

Clinicians also differ on the blood sugar values which are diagnostic of diabetes. This difference of opinion increases the problem of determining the most suitable value to use in screening for this condition.

Bondy (1955) has commented that "the diagnosis of diabetes cannot be discarded until a normal glucose tolerance test has been obtained, nor can it be considered established unless either a considerably elevated fasting blood glucose or an abnormal response to the glucose tolerance test has been found." However, the personal physicians to whom positive screenees are referred in the United States do not usually carry out this recommendation. McLoughlin et al. (1953) found that of 443 screenees diagnosed as having no disease, six had glucose tolerance tests and 162 had fasting blood sugar tests. The remaining 275 had been ruled out by urinalysis and "clinical judgement." It would appear that many physicians are either not aware of, or do not agree with Joslin's (1952) statement

that diabetes can be present in the absence of symptoms and clinical signs. Similarly inadequate diagnostic examinations have been shown to occur in Washington, D.C. (Loube and Alpert, 1954).

McLoughlin et al. (1953) suggest that instructions in the proper method of diagnosis of diabetes mellitus should be included in information sent to physicians when patients are referred. Family physicians are not likely to welcome such a step, however. An alternative plan may be for the screening agency, with the approval of the local physicians, to make the necessary diagnostic examination before the referral is carried out.

Fasting Blood Glucose Values

Blood glucose levels of 130 mg. per cent (Folin and Wu) or 110 mg. per cent (true glucose) are frequently regarded as diagnostic of diabetes in persons who have not eaten for twelve or more hours (Cantarow and Trumper, 1955). However, many mild diabetics have normal fasting values, and would be missed if fasting values were used in screening (Joslin, 1952; Cantarow and Trumper, 1955; Mosenthal and Barry, 1946).

Post Prandial Blood Glucose Values

Diabetics show a greater rise in post prandial blood sugar levels than non diabetic persons. It is wise to use higher screening levels for persons who have eaten within two hours than for those beyond two hours (Wilkerson et al.,1955; Kenny and Chute, 1953). Wilkerson et al.(1955) suggest that screening persons beyond two hours after eating is less efficient than

screening those within two hours post prandially.

The variation in quantity and quality of food eaten, and the unreliability of the time stated to have elapsed since eating will result in a considerable overlap in the post prandial values of diabetic and non diabetic persons. No matter what screening level is used, therefore, this procedure is likely to be less sensitive and less specific than giving a definite dose of sugar at a known interval before the test.

Blood Glucose Values after Sugar Ingestion

The oral glucose tolerance test is most frequently used in diagnosing diabetes. The height of the peak and the duration of the elevated tolerance curve are closely similar whether 50 or 100 gm. of glucose is used (Gray, 1923; Mosenthal and Barry, 1950). The smaller amount of 50 gm. will be just as satisfactory in screening for diabetes, and will be more acceptable to screenees.

Sucrose produces a similar tolerance curve to that following ingestion of the same weight of glucose (Gray, 1923). This test will be less acceptable to the screenee, however, since sucrose is one and one-third times sweeter than glucose. Moreover, physicians may less frequently accept the results of sucrose tolerance tests, since the similarity of results with glucose and sucrose is not well known to the medical profession.

The effect of glucose taken recently after food has not been well studied. Food will probably slow down the absorption of glucose, however, with little or no increase in the post prandial blood glucose values. Glucose ingested beyond two hours post prandially will produce a higher rise in blood

sugar than when ingested within two hours, since absorption will be less delayed by the food which has already left the stomach.

An additional dose of glucose will not unduly raise the blood sugar values of normal screenees, and will sufficiently elevate the values in diabetic patients who have not eaten previously that they will be above the screening level. For this reason, 50 gm. glucose was routinely given to each screenee on arrival at the Baltimore clinic.

The Glucose Drink

To prevent the need for weighing out glucose and making up individual drinks, a flavouring company and a carbonated beverage manufacturer helped to produce a bottled carbonated drink. An orange flavouring agent, with a certain amount of citric acid, was used to overcome the sweet taste of the glucose. The method of manufacture was approximately as follows:

The orange flavouring agent with citric acid, the powdered glucose and a 25 per cent solution of sodium benzoate were mixed and heated to produce a syrup. The quantities used were such that two ounces of the syrup contained 50 gm. glucose, and that the sodium benzoate produced a one-twentieth of one per cent solution in the final drink. Two ounces of the syrup were mechanically dropped into seven ounce bottles, to which five ounces of mildly carbonated water were added.

A small number of random samples of the final drink were tested for glucose content. The tested samples had a median glucose content of 48 gm., with an average error of two gm.

or four per cent around this value. The total cost of the drink, when delivered to Bennett Hall, was 0/6d (5.9 cents) per bottle. Less than one per cent of screenees disliked the drink, either because of its sweetness or its carbonation.

Wilkerson-Heftmann Blood Glucose Test.

The Wilkerson-Heftmann (1948) test is a modification of a true glucose method devised by Hagedorn and his co-workers (1946). The reagents are commercially available in tablet form, and restrict the choice of screening levels to three values -- 130, 160 and 180 mg. per cent.

The Hewson Clinitron is a device which automatically adds the reagent tablets to the pipetted blood and heats the constituents. The result is available five minutes after the Clinitron receives the tube containing the pipetted blood. The Clinitron has a maximum capacity of 120 estimations per hour. At this rate, however, more than one technician is needed to pipette blood, clean test tubes and pipettes, and to read and record results. In the Baltimore clinic a previously completely untrained person could comfortably carry out 25 estimations per hour after one day's experience.

Inadequate heating of the test tube with incomplete removal of the blood protein, failure of one or more of the reagent tablets to drop in, or inadequate cooling of the test tube at the end of the test will produce false positive results. Every positive blood specimen was routinely retested in the Baltimore clinic to detect these accidental false positives.

The Wilkerson-Heftmann test is stated to be accurate within ± 5 mg. per cent, although there are few published reports

to confirm this finding (Wilkerson and Heftmann, 1948). This figure may be attainable under ideal conditions, but is probably too small for the test as carried out under screening clinic conditions.

The rapidity of the test and its great saving in labour are the main advantages of the Wilkerson-Heftmann procedure. Other screening clinics have used the Folin and Wu, Somogyi-Nelson, anthrone, and picric acid methods, all of which require the presence of a trained technician and take twice as long.

Urine Sugar Tests used in Screening

Benedict's test is the most widely used in clinical laboratories, and has been used fairly commonly in screening. It has the disadvantages that liquid reagents and heating are necessary.

The Clinitest method uses a proprietary tablet (distributed by Ames Co.(London) Inc.) which does its own heating when added to a one-in-two dilutions of urine in water. It is believed to be quite accurate, but is less sensitive than Benedict's test (Olmstead et al.,1953).

The Galatest method uses a proprietary powder, to which a small drop of urine is added. The Dreypak technique is an adaptation of Benedict's test to make its mass application more practicable (Olmstead et al.,1953). The screenee dips a strip of absorbent paper into the urine specimen. When dry, and bearing information about the screenee, the strip is mailed to a central laboratory which carries out a modified Benedict's test.

Urine Sugar versus Blood Sugar Tests in Screening.

Urine sugar tests have the advantage that wide use is possible, without the need for trained technicians. However, since high and low renal thresholds will result in false negative and false positive results respectively, this method is less sensitive and less specific than blood sugar screening tests. A considerable number of studies confirm the superiority of the blood sugar tests.in screening for diabetes. In at least two programs, however, the urine and blood sugar tests appeared to give equally satisfactory results (U.S.Public Health Service, 1953; Weinerman et al.,1952).

Procedure in the Baltimore Clinic

After ingestion of glucose, diabetic persons have blood sugar values higher than usual after one and two hours. A small proportion of non diabetic persons have elevated one hour values but normal two hour values. The procedure used in the Baltimore clinic is based on this reasoning.

On arrival at the clinic the screenee drank the carbonated orange drink containing 50 gm. glucose. The urine sugar was tested 30-50 minutes later, a positive result being one plus or more by the Clinitest method.

The venous blood glucose level was estimated 45-75 minutes after the screenee had taken the drink. When the result was above 160 mg. per cent by the Wilkerson-Heftmann method the screenee was asked to remain to have a two hour specimen taken to be screened at 130 mg. per cent.

A small number returned at a later date to have the two

hour test carried out. A few were unable to have the two hour value taken at any time. Those individuals with glycosuria, but whose blood glucose was below 160 mg. per cent at about one hour, were not retested but considered to have a low renal threshold.for glucose.

Screenees with glycosuria, whose blood sugars were positive at one hour and were either confirmed positive or not done at two hours, were referred as probable diabetics. Those without glycosuria but with elevated one and two hour blood glucose values were advised to visit their physicians in order that diabetes could be more definitely confirmed or excluded. These latter persons were not followed up to obtain the diagnostic findings.

Screenees with glycosuria, but whose blood sugar level was below the screening level at either one or two hours, were told that they did not appear to have diabetes, and the results of the tests were explained to them. A few persons, especially when obese, probably visited their physicians for advice on whether any precautionary measures were advisable.

Results of the Baltimore Clinic.

Table 1 - Results of the Blood Sugar Screening Test for Diabetes, by Age and Sex of Screenees

Screening Results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YEARS				
		* Total	<30	30-44	45-64	65+	* Total	<30	30-44	45-64	65+
Number Screened	1916	914	171	403	283	55	999	229	387	300	72
Negative 1 hr.value	1856	887	169	395	270	51	966	226	378	285	66
Positive 1 hr.value	60	27	2	8	13	4	33	3	9	15	6
Retested	44	23	1	7	11	4	21	2	7	9	3
Positive 2 hr.value	17	10	1	0	6	3	7	0	2	3	1
1 hr. +, 2 hr. + or not obtd.	33	14	2	1	8	3	19	1	5	9	4

Footnotes to Table 1:
1/ Includes unknown age and sex. *Includes unknown age.

Table 1 of this chapter shows the results of the blood sugar screening procedure only. Seventeen screenees had both positive one and two hour blood sugars. An additional sixteen screenees had a positive one hour blood sugar, on whom a two hour value was not obtained. Of these 33 screenees, only those with glycosuria were referred as probable diabetics.

Of those screenees with positive one hour values and who were retested, 39 per cent had positive two hour values. The number of positives is therefore reduced by two-thirds by this retesting procedure. It is likely that most,if not all of the true positives are among those positive at two hours. (No diagnostic examination was carried out, however, on those with positive one hour and negative two hour results to determine if some diabetics were being lost by the retesting procedure).

Table 2 on the following page shows that, of the 54 screenees with glycosuria, 15 (28%) had positive blood sugar tests. The urine sugar test alone produces a considerable proportion of false positives in screening for diabetes.

Fourteen of the 33 screenees classified as positive by the blood sugar tests were referred as probable diabetics because they had glycosuria and were previously unaware of the condition. The remaining 19 were advised, but not urged to visit their personal physician. Follow-up reports were not obtained on this latter group. Of the nine screenees on whom the follow-up was completed, five were confirmed to have diabetes and four were previously unknown to the family physician.

Table 2 - Results of Screening for Diabetes Mellitus and of Follow-up with the Physician, by age and sex of Screenees

Screening and follow-up results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YEARS				
		* Total	<30	30-44	45-64	65+	* Total	<30	30-44	45-64	65+
Number screened	1916	914	171	403	283	55	999	229	387	300	72
Positive urine sugar	54	33	5	8	14	6	21	8	4	8	1
Blood S. + at 1 hr, + or not done at 2 hrs.	33	14	2	1	8	3	19	1	5	9	4
Both tests positive	15	9	1	0	6	2	6	0	1	5	0
Diabetes previously unknown	14	8	1	0	5	2	6	0	1	5	0
Follow-up with MD completed	9	5	1	0	3	1	4	0	0	4	0
Diabetes confirmed	5	3	0	0	2	1	2	0	0	2	0
Previously unknown to MD.	4	3	0	0	2	1	1	0	0	1	0
Rate per 1000 persons screened.											
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Positive urine sugar	28	36	29	20	49	109	21	35	11	27	14
Blood S. + at 1 hr, + or not done at 2 hr.	17	15	12	2	28	55	19	4	13	30	56
Both tests positive	8	10	6	0	21	36	6	0	3	17	0
Diabetes previously unknown	7	9	6	0	18	36	6	0	3	17	0
Follow-up with MD. completed	5	6	6	0	11	18	4	0	0	13	0
Diabetes confirmed	3	3	0	0	7	18	2	0	0	7	0
Previously unknown to MD.	2	3	0	0	7	18	1	0	0	3	0

1/ Includes unknown age and sex. *Includes unknown age

The results of screening for diabetes, by race and sex of the screenees, are shown in Table 3 on the following page.

Table 3 - Results of Screening for Diabetes Mellitus, by Race and Sex of Persons screened.

Screening results	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Number screened	1916	1484	728	756	412	175	237
	Rate per 1000 persons screened						
All results	1000	1000	1000	1000	1000	1000	1000
Normal	962	965	957	972	952	960	945
Abnormal urine sugar	28	26	38	15	36	29	42
Abnormal blood sugar	17	16	14	19	22	23	21
Both tests abnormal	8	7	10	5	10	11	8

1/ Includes unknown race and sex. Excludes persons on whom only one test was completed.

Discussion

Two previously unknown diabetics per 1000 screenees is a low yield, comparing quite unfavourably with the study of Wilkerson and Krall (1947) in Oxford, Massachusetts. These investigators suggest that the prevalence of previously unknown diabetes in the population is about 14 per 1000.

Several factors played a part in producing the low yield in Baltimore. The required presence of glycosuria in the referrals probably prevented the referral of a number of true diabetics, causing the test to be insensitive. A number of referrals did not visit their family physicians. The diagnostic examinations were carried out usually by the family physicians. It is possible that some of these examinations were inadequate, and that some true diabetics were missed.

All 14 referrals had true blood glucose values above 160 mg. per cent at one hour, with glycosuria. About half had two hour blood sugar values above 130 mg. per cent. It is likely, therefore, that almost all were diabetics. However, only 56 per cent of those on whom the follow-up was completed were confirmed to have diabetes.

As shown in chapter 5 on obesity, a large proportion of the positive tests occur in the non obese screenees. This finding confirms other studies (Futcher and Marcus, 1956) which show that diabetes screening should not be confined to overweight subjects.

Conclusions and Summary

A modified glucose tolerance test was used to screen for diabetes mellitus. The yield of newly discovered diabetics was low compared with other studies. This low yield was believed due partly to the failure to refer screenees with positive blood sugar tests without glycosuria, and partly to some inadequate diagnostic examinations after referral.

These defects may be overcome by omitting the urine sugar test in screening, and by providing adequate diagnostic facilities for the positive screenee.

REFERENCES

Allen,F.M. in Foreward to Ref. Joslin et al. (1952).

Bodansky,O. (1952). Biochemistry of Disease. The Macmillan Co., P 500-501.

Bondy,P.K. (1955). 1954-55 Year Book of Medicine. The Year Book Publishers Inc.

Bowman,W.M. and Enterline,P.E. (1954). Public Health Rep.69:240

Cantarow,A., and Trumper,M. (1955). Clinical Biochemistry. Philadelphia and London. W.B.Saunders Co. 5th Ed.

Chesrow,E.J., and Bleyer,J.M. (1954). Geriatrics 9:276.

Collens,W.S. (1954). J.Clin.Nutr. 2:195.

Copp,E.F., and Barclay,A.J. (1923). J.Metab.Research 4:445.*

Deren,M.D. (1937). J.Lab.Clin.Med. 22:1138.

Ditzel,J., and Sagild,U. (1954). New England J.Med.250:541.

Dolger,H. (1947). J.Amer.Med.Ass. 134:1289.

Duncan,G.G. (1951). Diabetes Mellitus. W.B.Saunders Co.

Dunlop,D.M. (1954). Brit.Med.J., Aug.14, P383.

Folin,O., and Wu,H. J.Biol.Chem. 41:367, (1920)*

Futcher,P.H., and Marcus,D.M. (1956). J.Chronic Dis. 3:294.

Gray,H. (1923). Arch.Int.Med. 31:241.

Hagedorn,H.C. et al. (1946). Swift Methods for Determination of Blood Sugar by Means of Potassium Ferricyanide. Reports of the Steno Memorial Hospital and Nordisk Insulinlaboratorium, Vol. 1. *

Haunz,E.A., and Keranen,M.T. (1950). Proc.Am.Diab.Ass. 10:200.

Jackson,W.P.U. (1955). The Lancet 2:625.

Joslin,E.P. et al. (1952). The Treatment of Diabetes Mellitus. Lea and Febiger, 9th Ed.

*Original article not read.

Kenny,A.J., and Chute,A.L. (1953). Diabetes 2:187.

Loughlin,W.C. et al. J.Lab.Clin.Med. 28:1165, (1943).

Lukens,F.D.W. et al. Endocrinology 32:475, (1943).*

Lukens,F.D.W., and Dohan,F.C. (1940). Science 92:222.*

Mattison,B.F. (1955). Public Health Rep. 70:625.

McLoughlin,C.J. et al. (1953). J.Amer.Med.Ass. 153:182.

Mosenthal,H.O. (1947). Med.Clin.N.Amer. 31:299.

Mosenthal,H.O., and Barry,E. (1946). Am.J.Digest.Dis. 13:160.

Mosenthal,H.O., and Barry,E. (1950). Ann.Int.Med. 33:1175.

Newburgh,L.H., and Conn,J.W. (1939). J.Amer.Med.Ass. 12:7.

Olmstead,W.H. et al. (1953). Diabetes 2:37.

Parkhurst,L.W., and Betsch,W.F. (1955). Med.Clin.N.Amer.39:1571.

Pincus,G., and White,P. in Ref. Joslin et al. (1952).

Richardson,G.O. (1953). Diabetes 2:454.

Rickets,H.T. (1954). Diabetes 3:80.

Sander,F.W. (1923). J.Biol.Chem. 58:1.*

Sunderman,F.W. et al. (1951). Am.J.Clin.Path. 21:901.

Todd,J.C. et al. (1953). Clinical Diagnosis by Laboratory Methods. The W.B.Saunders Co. 12th Ed.

Tolstoi,E. (1953). The Practical Management of Diabetes. Charles C. Thomas, Publisher.

Tolstoi,E., and Weber,F.C.,Jr. (1939). Arch.Int.Med. 64:91.

Tolstoi,E. (1950). Med.Clin.N.Amer. 34:485.

U.S.Public Health Service (1953), Division of Chronic Disease and Tuberculosis. A Summary of Some Diabetes Screening Projects. P 4-5.

Wagner,A. (1955). Geriatrics 10:588.

*Article not read in the original.

Weinerman,E.R., et al. (1952). Am.J.Pub.Health 42:1552.

Wilder,R.M. (1948). J.Amer.Med.Ass. 138:349.

Wilkerson,H.L.C. (1949). New York State J.Med. 49:2945.

Wilkerson,H.L.C.,et al. (1955). J.Chronic Dis. 2:461.

Wilkerson,H.L.C. and Heftmann,E. (1948). J.Lab.Clin.Med. 33:236.

Wilkerson,H.L.C. and Krall,L.P. (1947). J.Amer.Med.Ass. 135:209.

Wolfson,W.Q. (1954). Glandular Physiology and Therapy. J.B.Lippincott Co.,5th Ed.

General References.

Bell,G.H.,Davidson,J.N.,and Scarborough,H. (1953). Textbook of Physiology and Biochemistry. The Williams and Wilkins Co.

Best,C.G., and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. The Williams and Wilkins Co.,10th Ed.

Conybeare,J., and Mann,W.N. (1952). Textbook of Medicine. The Williams and Wilkins Co.

Garland,H.G., and Phillips,W. (1953). Medicine. MacMillan and Co.,Ltd.

Harrison,T.R.,Editor (1954). Principles of Internal Medicine. The Blackiston Div.,McGraw-Hill Book Co.,2nd Ed.

Price,F.W. (1952). A Textbook of the Practice of Medicine. Oxford University Press, 9th Ed.

Sunderman,F.W., and Boerner,F.,(1950). Normal Values in Clinical Medicine. W.B.Saunders Co. 1st Ed.

Wilkerson,H.L.C., and Glenn,B. (1951). A Preliminary Report on a Study of Blood Sugar Screening Levels. Am.J.Pub.Health 41:440.

Wright,S. (1952). Applied Physiology. Oxford University Press, 9th Ed.

SCREENING FOR HYPERTENSION

Section C
Chapter 2

SCREENING FOR HYPERTENSION

Hypertension is a sustained, abnormally high level of systolic or diastolic blood pressures or both. When all disorders known to cause secondary hypertension are absent, the condition is named essential or primary hypertension.

Clinicians do not agree on the arbitrary criteria used for the upper limit of normal pressure. Probably the most commonly used are 140/90 and 150/90 mm. of mercury (Likoff and Davie, 1953; Burch, 1953; Coneybeare and Mann, 1952). Other clinicians have proposed higher levels such as 160/100 (East and Brain, 1948), 180/100 (Perera and Atchley, 1947) and even 180/110 mm. mercury (Evans, 1948).

While it is usual to employ only one level as the abnormal value for all ages and both sexes, a few studies have differed. Luisada (1954) suggests 120/75 for adult males and 113/70 for women. Master et al. (1952) recommend a sliding scale of values, going up even higher than any so far mentioned.

Perera (1954) has given clinical evidence, which requires further documentation, that hypertensive vascular disease can progress even when a normal blood pressure is present. The exact blood pressure level may therefore have little clinical significance. Perera comments that "we must ask ourselves whether the mere lowering of blood pressure, by whatsoever means, is a sufficient therapeutic objective."

The use of high blood pressure levels in screening, therefore, may not necessarily find cases of hypertension at an advanced stage. Even the use of low pressure screening levels may not detect all cases of hypertensive disease, since hyper-

tensive changes are said to occur in the presence of a normal blood pressure.

On the other hand, the family physician to whom the screenee is referred will usually use the blood pressure level as his main criterion for diagnosis, since there is no adequate test for hypertensive vascular disease in the absence of hypertension. The higher the blood pressure level of the screenee, the more likely he is to be diagnosed as hypertensive by his personal physician.

Labile Blood Pressures

Most clinicians may support the belief that a labile blood pressure is a pre-hypertensive sign; that a pressure which rises temporarily may eventually develop into a permanently raised blood pressure.

For some years the cold pressor test has been used to measure lability of blood pressure. This test measures the blood pressure rise produced by immersing one hand in ice water. It is fairly well established that the changes are more marked in persons who are already hypertensive. At present, studies are in progress to determine whether persons who react strongly to this test are likely to develop hypertension in the future. Fishberg (1954) states, however, that the test is not a dependable one, and is certainly not sufficiently reliable for use in screening pre-hypertensive persons.

Perera (1955) has found that persons with a non labile hypertension have a briefer duration of hypertension and lower average age at death than those with a labile blood pressure.

He believes that the determination of "resting" blood pressure, obtained under basal conditions, is important in evaluating the hypertensive patient. He suggests that casual values, taken irrespective of whether the person has rested or not, have little prognostic significance. These findings are based on a retrospective study of fifty selected dead patients, and the validity of the conclusions is difficult to determine.

However, in screening for hypertension it does seem reasonable to take the pressures under conditions producing as little physical and mental stress as possible. Clinicians who believe that help can be given to pre-hypertensive patients, who might have normal results under such conditions, may disagree with this suggestion.

Symptoms of Hypertension

There is now considerable doubt that there are any specific symptoms of hypertension. It has been noticed, for instance, that the so-called typical hypertensive symptoms occur much more commonly in persons who know that their blood pressure is elevated, and that these symptoms are frequently relieved by therapy which has no effect on the blood pressure.

Conybeare and Mann (1952) comment "..It is indeed doubtful whether a high blood pressure per se produces any symptoms." Semisch and Lewis (1954) emphasise that "fluoroscopic evidence of left ventricular enlargement or electrocardiographic evidence of hypertrophy...usually precedes symptoms..."

It would therefore appear that a questionary including the alleged symptoms of hypertension is not likely to be helpful in screening for hypertension.

Hypertension and Cardiac Enlargement

The chest x-ray does not detect the gradual left ventricular hypertrophy which occurs with persistent hypertension. Detectable x-ray changes occur only in the later stages of the disease when cardiac dilatation takes place. It is not likely, therefore, that chest x-rays of apparently well persons will detect cases of hypertension which are missed by the blood pressure reading.

Hypertension and the Electrocardiogram

The left ventricular hypertrophy, which occurs at an earlier stage in the hypertensive process than cardiac dilatation, will produce electrocardiographic changes. Bourne (1955) suggests that "if the electrocardiogram shows no evidence of left ventricular hypertrophy, and normal QRST complexes, it is unlikely that serious hypertension is present."

It is therefore possible that the electrocardiogram may help to indicate those hypertensives who are most in need of treatment. Little information is available, however, on whether the screening electrocardiogram is likely to detect cases of hypertension which are missed by the blood pressure reading.

Indirect Method of Determining Blood Pressure

The sphygmomanometric method of determining blood pressure is the most widely used. As originally described by Korotkow in 1905, the brachial sound goes through four phases. Phase IV begins when the sound becomes muffled in quality, and ends when the sound disappears.

The significance of phase IV has produced much contro-

versy over reading the diastolic pressure. The point where muffling occurs is most commonly recommended and used in Great Britain as the diastolic pressure. A comparison of indirect readings taken at this point and simultaneous measurements of arterial pressure obtained by direct intra-arterial manometry shows that the muffling occurs nine mm. mercury above the true diastolic pressure (Hamilton et al.,1936).

The 1951 recommendations of a United States Committee to Revise Standardization of High Blood Pressure Readings (American Heart Association, 1952) advise that "the point of complete cessation is the best index of diastolic pressure." This recommendation was carried out in the Baltimore clinic.

The difference between the point of muffling and the point of disappearance has been variously stated to be between 4 and 14 mm. mercury. The use of the higher value (beginning of phase IV) will result in the referral of a significantly larger proportion of possible hypertensives from a screening clinic. It is therefore important to mention the method of reading diastolic pressure in reporting the screening level and results.

Effect of Position on Blood Pressure

Schneider and Truesdale (1922) found that the mean systolic and diastolic pressures increased slightly on standing. However, only a proportion of their subjects showed this increase, while a decrease occurred in others.

After studying this and other reports, the Committee to Revise Standardization of High Blood Pressure Readings

(American Heart Association, 1952) concluded that there is no significant difference in the blood pressure taken in the lying or sitting position. This opinion is not universal, however, as can be seen by Bourne's (1955) recommendation that "the routine blood pressure should always be taken with the patient lying flat, for in some individuals it will be lower in the standing and in the sitting than in the recumbent position."

In the Baltimore clinic, the blood pressure was taken after the screenee had been lying down for three or four minutes.

Inequality of Blood Pressure in Two Arms.

Studying the results of simultaneously taking the blood pressure on the two arms, Korns and Guinand (1933) found that higher pressures occur much more commonly on the right side. Of 1000 apparently normal persons 222 had a systolic inequality and 217 had a diastolic inequality of ten mm. of mercury or more. This inequality was often found to be transitory, and might involve the systolic only, the diastolic only, or both.

The significance of this finding is uncertain, since it is known that a mean error of $\pm$ 8 mm. mercury may be expected in individual readings of systolic and diastolic pressures (American Heart Association, 1952). The apparent inequality might at least have been partly due to this error in reading.

The available evidence does not suggest that all readings need be taken on one particular side in screening for hypertension.

Hypertension and Obesity

The initial findings of the Albany Cardiovascular Health

Center (James and Hilleboe, 1955; James et al.,1955) suggest that in each age group the proportion with elevated blood pressure in persons over 20 per cent above the Metropolitan Life Insurance Standards for height and weight is about twice that in the non obese group. This result confirms the findings of other studies (Thomas and Cohen, 1955).

Preble (1923) studied 194 obese persons before and after reducing their weight by ten pounds or more. There was an average drop in blood pressure of 18 mm. systolic and 10 mm. diastolic. Twenty-two persons with a systolic pressure of 200 mm. mercury or more underwent a similar reduction in weight. The average systolic drop was 43 mm., while the average diastolic fall was 21 mm. mercury. Studies by Spencer (1929) and John (1929) confirm that the blood pressures of obese hypertensives tend to fall with weight reduction. Wheatley (1955) suggests that weight reduction of moderately overweight men will reduce their mortality from 142 per cent of the standard to 113 per cent. However, Pickering (1955) believes that it is quite uncertain whether reduction in weight will improve the prognosis in obese hypertensives.

Ragan and Bordley (1941) found that the unusually small arm may result in indirect systolic readings 30 mm. mercury lower than the actual intra-arterial systolic pressure. An unusually large arm can produce an indirect reading which is too high by the same amount. They therefore recommended that "statistical studies of the relation between blood pressure and body weight should take into account the influence of the circumference of the arm upon the accuracy of the blood pres-

sure measurements." Days and Scrimshaw (1953) support this conclusion.

Pickering et al. (1954) analysed the data of Ragan and Bordley. They commented that "variations in arm circumference do not account for more than a quarter of the variation of differences between direct and indirect readings. Persons of the same arm circumference may show widely divergent differences and so it is doubtful how far it is useful to apply a correction for arm circumference to an individual reading; the correction will often make the discrepancy between auscultatory measurement and true arterial pressure worse instead of better...Arm circumference in adults is but little related to age and hence the pattern of increase of arterial pressure with age turns out not to be significantly altered by applying the correction."

In summary, there appears to be no need to correct for arm circumference in screening for hypertension. Although arm circumference may affect the indirect blood pressure readings, the evidence available is not sufficient for corrections to be made for this factor in epidemiological studies of hypertension.

Normal Blood Pressure and Sex

Master et al. (1952) found that male average systolic pressures were greater than the average values for women in the age groups under 45 years, and that the female average pressures were greater than the corresponding male values over

45 years. The same findings applied to the diastolic pressures except that the age for the change-over was 50 years. A previous study by Symonds (1923) showed similar differences between the two sexes.

A recent study of hospital outpatients by Hamilton et al. (1954) shows average systolic pressures in men as higher only until 35 years, above which the values in women are consistently higher. The diastolic values for women are almost always higher than the corresponding values for men. This study, therefore, does not agree completely with those mentioned above.

In practice the sex difference is sufficiently small to permit using the same screening levels for both sexes. The figures of Master et al.(1952), for instance, show a difference of 5 to 10 mm. mercury for the upper limit of normal for systolic pressure, and a difference of 1 to 4 mm. mercury for the upper limit of normal for diastolic pressure between the two sexes.

Blood Pressure Changes with Age

"A systolic pressure of 150 is to be regarded as abnormal at all ages...A diastolic pressure of 100 is probably always pathological, and one of 95 mm. very suspicious." So states Fishberg (1954) in his textbook Hypertension and Nephritis.

Such a view has strong support in medical literature. "It is now practically universally believed that the blood-pressure does not increase with advancing age as much as was previously thought to be the case" (Price, 1950). "A pressure of 115 mm. is just as normal, and a pressure of 140 is just as abnormal in an old man as in a young man" (Alvarez and Stanley, 1939). Lechgaard (1946) placed the upper limit of

normal pressure at 140/90 at all ages. Fife et al.(1954) used a diastolic pressure of 92 mm. mercury or above to indicate hypertension at all ages.

Robinson and Brucer (1939) studied over 10,000 insurance policy holders and concluded that a blood pressure over 120/80 is abnormal at any age and indicates incipient hypertension. Since this study is widely quoted, it might be helpful to examine the findings in detail.

The majority of Robinson and Brucer's group came in for a routine periodic examination given free by the insurance company. All were accepted risks, although most persons came into the study long after the original insurance examination. They usually had no complaints or only minor ones at the time of examination.

Robinson and Brucer explain:"One must devise some method of roughly separating the pathologic from the normal pressures before the group is subjected to an extensive statistical analysis...During the last ten years an ever increasing number of authorities in this field have placed 140 mm. as the upper limit of normal systolic pressure...There seems to be an equal consensus that the upper limit of normal diastolic pressure should be placed at 90 mm." Robinson and Brucer therefore discard from their study all persons who have pressures above these levels. Thus, from each succeeding age group an increasing proportion of subjects is removed on the basis that they had hypertension and would "distort the picture of the normal blood pressure averages." They comment that the exclusion "does not

alter the conclusions of the analysis."

The authors therefore begin their study with the preformed conclusion that 140 mm. systolic and 90 mm. diastolic is abnormal at any age. The arbitrary removal of such persons from a study, the purpose of which is to determine the upper limits of normal pressure, can hardly fail to alter the conclusions of the analysis.

Examining each ascending age group, Robinson and Brucer find that around 25 per cent of persons in the deliminated sample had systolic pressures under 110 mm. They reason that "apparently..most men who have low pressures (under 110 mm.) do not show any rise in systolic pressure throughout their lives.."

The investigators overlooked the fact, however, that they had removed an increasing number of persons in the "above 110 mm." group with each increase in age. To keep the proportion of persons below 110 mm. at 25 per cent, a number of this group would have to develop higher pressures as age increased.

Using similar methods of reasoning, Robinson and Brucer conclude that "..about 40% of the adult population is prehypertensive or hypertensive...The normal range of systolic pressure for men and women is from 90-120 mm. Hg. and diastolic pressure 60-90 mm. Hg." The numerous pitfalls and the erroneous interpretations in this study suggest, however, that the conclusions are not valid.

The literature to support the belief that "normal" pressures rise with advancing age is also abundant. "One pair of

figures for the systolic and diastolic pressure cannot apply to all age groups and both sexes.." (Master et al.,1952). "There is common agreement that average diastolic and systolic pressures rise with advancing age" (Shock, 1952). Morsell (1951) recommends that the "normal range" should rise with increasing age. Russek (1948) takes an intermediate stand by suggesting that "normal standards" should "allow a progressively increasing range in the systolic blood pressure while permitting but little variation in the diastolic level with succeeding decades."

Master, Garfield and Walters (1952) sampled the records of 74,000 persons in industrial plants and U.S.Army airfields to produce their extensive analysis of blood pressure readings. The majority of the persons studied were at work, although some were those who applied for work and who may have been rejected. Their only manipulation of the records was the rejection of those taken in areas where the blood pressure readings were not considered to have been accurately read.

The investigators found that the mean systolic and diastolic pressures increased steadily throughout life. They arbitrarily decided to call as normal those persons in each five year age group who were within 1.282 standard deviations of the mean; those between this point and 2 standard deviations were placed in the borderline group, who were probably normal but required study before the final diagnosis was established; those beyond 2 standard deviations were probably hypertensives, although final clinical confirmation was required.

Since the distribution curves for each age group are not typically "normal" or symmetrical, the use of the mean and

standard deviation for dividing the groups is subject to criticism. In practice, however, the study seems to provide the clinician with a reasonable series of figures for each age group. Figure 1 of this chapter, on page 98, compares the screening levels used in the Baltimore clinic with the figures published by Master et al.

Even more recently Hamilton et al. (1954) studied all patients attending certain clinics on specific days in 1951. No patient was omitted for any cause whatsoever. Hamilton et al. found a greater rise with age of the mean systolic and diastolic pressures than Master and his group. Hamilton et al. suggest, as a possible reason for this difference, that middle aged persons with symptoms of hypertension do not continue to work in industrial plants. Perhaps an equally reasonable explanation may be that, just as hospital outpatients have a greater prevalence of tuberculosis than the general population, so they may have a higher prevalence of hypertension. The figures proposed by Master et al. may be more acceptable to clinicians, at least in the United States.

In summary, there is no valid evidence that an individual whose pressure remains below a certain level in early adult life will usually continue to have this low pressure throughout life. In screening for hypertension, therefore, it is recommended that the screening levels should be raised for persons in the older age groups.

The Need for Early Detection

More than two decades ago Burn (1934) asked: "Does the

Relation of screening levels to upper limits of "normal" blood pressure and lower limits of "hypertension" (averaged for both sexes). Values from Ref.12 of this chapter.

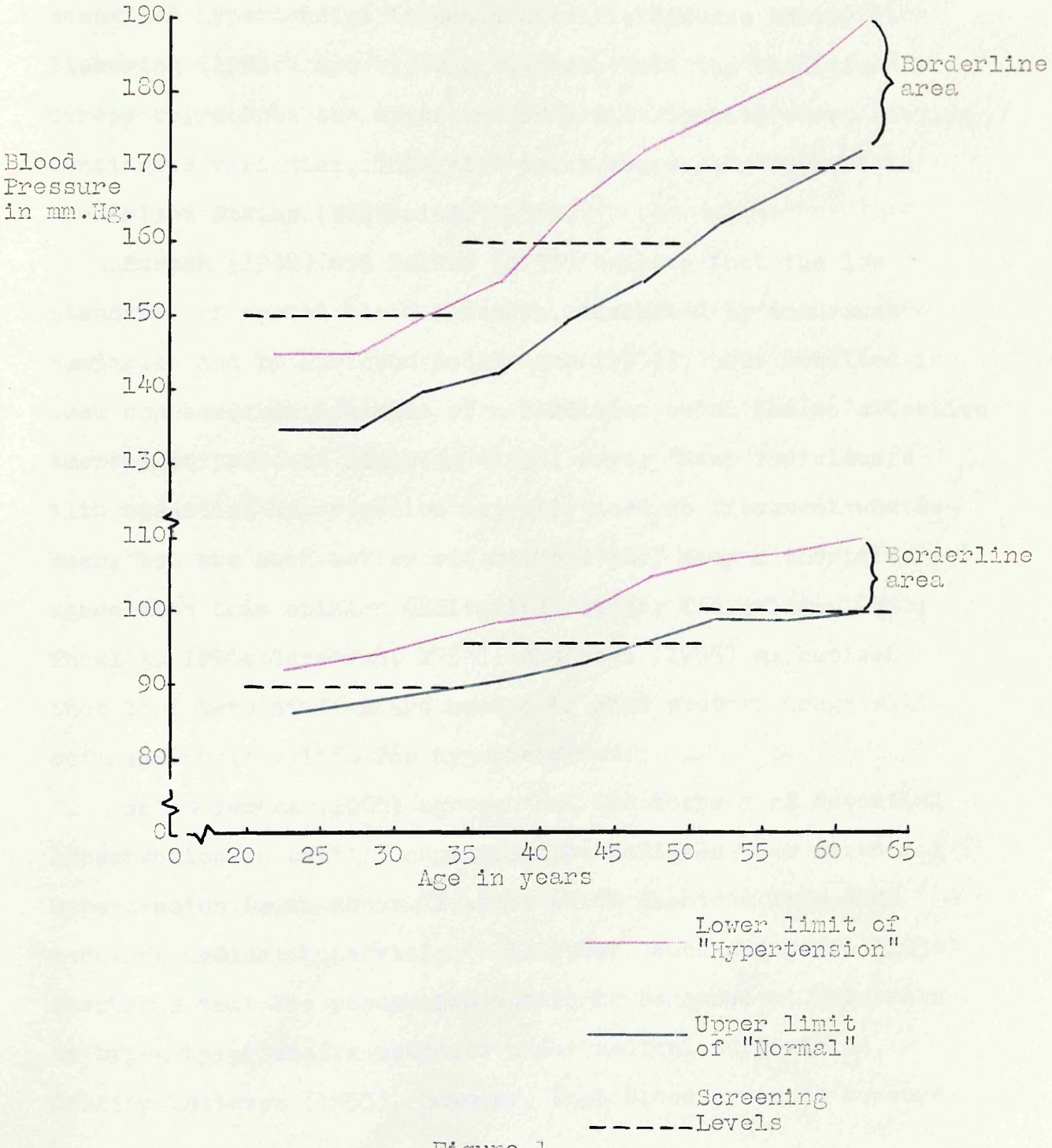


Figure 1.

disease known as essential hypertension really exist, or is it not a figment of the unrestricted imagination ?" These doubts have continued and even increased in recent years (Editorial, 1954a). Foremost among clinicians who believe that essential hypertension is not a specific disease entity is Pickering (1955 a and b), who suggests that the condition merely represents the upper end of a distribution curve showing continuous variation. This view gains increasing support in the United States (Editorial, 1956).

Russek (1948) and Palmer (1955) believe that the low standards of normal blood pressure, advocated by insurance companies and by Robinson and Brucer (1939), have resulted in much unnecessary treatment of a condition which has no effective therapy at present. Fishberg (1954) says: "Many individuals with essential hypertension not only need no treatment whatsoever, but are much better off without it." Many authorities agree with this opinion (Editorial, 1954b; Pickering, 1955b; Thewlis, 1954; Corcoran, 1954). Wheatley (1955) emphasises that long-term studies are needed to show whether drugs will actually prolong life for hypertensives.

While Perera (1956) agrees that the therapy of essential hypertension is still nonspecific, he believes that essential hypertension is an abnormal state which shortens life and requires medical supervision. Hilleboe, James and Doyle (1954) recommend that the population should be screened at intervals to bring hypertensive subjects under medical supervision. Getting believes (1955), however, that blood pressure measure-

ment is an unreliable and unnecessary screening procedure.

Other Screening Programs

The Richmond, Virginia, multitest clinic held in 1950 (American Medical Association, 1955) used a blood pressure screening level of 150/100 mm. mercury. 14.6 per cent of the screenees were above either of these levels. Ofthe total screenees, 3.5 per cent were confirmed to have hypertension when they visited their physicians, and 0.4 per cent were previously unknown. Thus, 5,451 positive screenees resulted in the known discovery of 162 new cases.

The San Francisco, California, survey of dock labourers (Weinerman et al.,1952) used a screening level of 170/95 mm. mercury. 21.0 per cent were found to be above either of these levels. 9.2 per cent of the total screenees were confirmed to have hypertension; and 5.2 per cent were new cases. The dock labourers had a considerable proportion of overweight persons, and probably produced a higher yield of hypertensives than would the general population.

Kurlander, Hill and Enterline (1955) studied hospital outpatients who had their blood pressure taken before going on to have a thorough clinical examination for cardiovascular disease. These outpatients had sought medical attention because of symptoms, not necessarily related to heart disease. The cases detected on screening were, therefore, likely to be further advanced in their disease than those detected by screening apparently well persons.

The screening test was regarded as positive if the blood pressure were above 150 mm. mercury systolic or 90 mm. mercury diastolic. This test detected 83 per cent of individuals with arteriosclerotic heart disease, 97 per cent of those with hypertensive heart disease, and 86 per cent of persons diagnosed as having heart disease without hypertension. Kurlander et al. concluded that the blood pressure appeared to be effective in detecting cases of hypertension with or without heart involvement, and cases of arteriosclerotic heart disease.

Procedure in the Baltimore Clinic

The blood pressure was taken, usually but not invariably on the right arm, after the screenee had been reclining for three or four minutes. The point of disappearance of the sound was used to determine the diastolic reading, except where the sound failed to disappear when the point of muffling was used. Only mercury sphygmomanometers were used.

The screening levels used were as follows:

Age below 35 years	150/90 mm. mercury.
Age 35 to 49 years	160/96 mm. mercury.
Age 50 or more years	170/100 mm. mercury.

The individual was referred for further examination when both systolic and diastolic pressures were higher than the screening level. When the systolic pressure alone was elevated, and was 20 mm. mercury or more above the screening level, or when the diastolic pressure alone was 10 mm. mercury or more above the screening level, referral was also carried out.

Results of the Baltimore Clinic

The screening and follow-up results are shown in Table 1 of this chapter.

Table 1 - Results of Screening for Hypertension and of Follow-up with the Physician, by Age and by Sex of Persons Screened.

Screening and follow-up results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YEARS				
		* Total	<30	30-44	45-64	65+	* Total	<30	30-44	45-64	65+
Number screened	2021	923	173	404	287	57	1095	265	432	313	74
Negative	1871	857	171	382	251	51	1011	264	416	264	58
Positive	150	66	2	22	36	6	84	1	16	49	16
Hypertension previously unknown	67	32	1	13	15	3	35	0	9	20	5
Follow-up with MD completed	41	20	0	9	9	2	21	0	6	12	3
Hypertension confirmed	31	13	0	6	6	1	18	0	4	11	3
Previously unknown to MD.	11	5	0	3	2	0	6	0	2	3	1
Rate per 1000 persons screened											
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Negative	926	929	988	946	875	895	923	996	963	844	784
Positive	74	71	12	54	125	105	77	4	37	156	216
Hypertension previously unknown	33	35	6	32	52	53	32	0	21	64	68
Follow-up with MD. completed	20	22	0	22	31	35	19	0	14	38	41
Hypertension confirmed	15	14	0	15	21	18	16	0	9	35	41
Previously unknown to MD.	5	5	0	7	7	0	6	0	5	10	14

1/ Includes unknown age and sex. *Includes unknown age.

Age. With the exception of males 65 and over, the proportion of positive screenees increases with age. This increase occurred

although higher screening levels were used for older age groups

Follow-up results. The follow-up with the physician was completed on 61 per cent of the positive screenees who were previously unaware of the presence of hypertension. 76 per cent of those screenees on whom the follow-up was completed were confirmed to have hypertension, but only 27 per cent were stated by the physician to be previously unknown cases.

False positive:True positive ratio. This ratio is satisfactory, being 1:3 when the figure for confirmed hypertensives is used. It is possible, however, that the unusually high screening levels resulted in some hypertensives being missed.

Table 2 - Results of Screening for Hypertension, by Race and Sex of Persons Screened.

Screening Result	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	2021	1569	735	834	432	177	255
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Negative	926	941	942	940	873	876	871
Positive	74	59	58	60	127	124	129

1/ Includes unknown race and sex.

Race. Table 2 of this chapter shows that the proportion of positive results is higher in non-white males and females than in the white groups.

Obesity. The screening results confirmed the frequently stated relationship between obesity and elevated blood pressure. This relationship is discussed more fully in Chapter 5 on Obesity.

Conclusions

In screening for hypertension, the use of higher screening levels for older age groups appears to be a rational and satisfactory procedure. The screening levels used in the Baltimore clinic resulted in a test of adequate specificity, but of unknown sensitivity.

Five previously unknown cases of hypertension were discovered per 1000 screenees. It is not known to what extent these cases benefit by their early discovery. There is much published literature to suggest that early discovery in the asymptomatic stage of hypertension is of little or no benefit.

REFERENCES

Alvarez,W.C., and Stanley,L. (1939). Arch.Int.Med. 46:17.

American Heart Association (1952). Recommendations for Human Blood Pressure Determinations by Sphygmomanometers. New York.

American Medical Association (1955). A Study of Multiple Screening. Council on Medical Service.

Bays,R.P.,and Scrimshaw,N.S. (1953). Circulation 8:655.

Bechgaard,P. (1946). Acta Med.Scandinav.,Supp.172,3.

Bourne,G. (1955). The Practitioner 174:129.

Burch,J.H. (1953). A Primer of Cardiology. Lea and Febiger, Philadelphia.

Burn,J.H. (1934). Proc.Roy.Soc.Med. 28:15.

Conybeare,J., and Mann,W.N. (1952). Textbook of Medicine. The Williams and Wilkins Co., Baltimore.

Corcoran,A.C. (1954). J.Iowa State Med.Soc. 64:457.

East,T., and Brain,C. (1948). Recent Advances in Cardiology. J. and A. Churchill Ltd. 4th Ed.

Editorial (1954a). The Lancet 1:1224.

Editorial (1954b). The Lancet 2:741.

Editorial (1956). Essential Hypertension. J.Amer.Med.Ass.160:673

Evans,W. (1948). Cardiology. Butterworth and Co.

Fife,R.,Marshall,E.A.,and Wright,J.H. (1954). Glasgow Med.J. 35:279.

Fishberg,A.M. (1954). Med.Clin.North Amer. 38:3.

Fishberg,A.M. (1954). Hypertension and Nephritis. Lea and Febiger, Philadelphia.

Getting,V.A. (1955). Quoted in A Study of Multiple Screening, P 60. Council on Medical Service, American Medical Association.

Hamilton,M. et al. (1954). Clin.Sci. 13:11.

Hamilton,W.F. et al. (1936). J.Amer.Med.Ass. 107:835.

Hilleboe,H.E., James,G., and Doyle,J.D. (1954). Amer.J.Public Health 44:851.

James,G. et al. (1955). New York State J.Med. 55:774.

James,G.,and Hilleboe,H.E. (1955). Amer.J.Public Health 45:140.

John,H.J. (1929). Endocrinology 13:388.*

Korns,H.M.,and Guinand,P.H. (1933). J.Clin.Invest. 12:143.

Kurlander,A.B., Hill,E.H., and Enterline,P.E. (1955). J.Chronic Dis. 2:427.

Likoff,W., and Davie,J.H. (1953). Clinical Cardiology. F.C.Massey, Editor. The Williams and Wilkins Co.,Baltimore.

1/ Master,A.H.,Garfield,C.I.,and Walters,M.B.(1952). Normal Blood Pressure and Hypertension. Lea and Febiger, Philadelphia.

Morsell,J.A. (1951). A Symposium on Essential Hypertension. The Commonwealth of Massachusetts.

Palmer,R.S. (1955). New England J.Med. 252:940.

Perera,G.A. (1954). Ciba Foundation Symposium on Hypertension. Little,Brown and Co.,Boston.

Perera,G.A. (1956). Circulation 13:321.

Perera,G.A. (1955). J.Chronic Dis. 1:121.

Perera,G.A.,and Atchley,D.W.(1947). Hypertensive Vascular Disease. Nelson Loose-Leaf Medicine, vol.IV. Thos.Nelson and Sons.

Pickering,G.W. (1955a). High Blood Pressure. Grune and Stratton, New York.

Pickering,G.W. (1955b). Ann.Int.Med. 43:1153.

Pickering,G.W., Fraser Roberts,J.W., and Sowry,G.S.C. (1954). Clin.Sci. 13:267.

Preble,W.E. (1923). Boston Med.Surg.J. 18:617.*

Price,F.W. (1950). Textbook of the Practice of Medicine. Oxford University Press.

Ragan,C.,and Bordley,J.(1941). Bull.Johns Hopkins Hosp.69:504.

Robinson,S.C., and Brucer,M. (1939). Arch.Int.Med. 64:409.

Russek,H.I. (1948). New York State J.Med. 48:1485.

Schneider,E.C., and Truesdell,D. (1922). Am.J.Physiol.61:429.*

1/ Luisada,A.A. (1954). Heart. The Williams and Wilkins Co., Baltimore.

*Original articles not read.

Semisch III,C.W., and Lewis,D.W. (1954). Med.Clin.N.Amer. 38:1730

Shock,N.W. (1952) in Cowdry's Problems of Ageing. A.I.Lansing, Editor. The Williams and Wilkins Co.,Baltimore.

Spencer,H.J. (1929). Med.Clin.N.Amer.12:604.*

Symonds,B. (1923). J.Amer.Med.Ass. 80:232.

Thewlis,M.W. (1954). The Care of the Aged. The C.V.Mosby Co., St.Louis.

Thomas,C.B., and Cohen,B.H. (1955). Ann.Int.Med. 42:90.

Weinerman,E.R. et al. (1952). Am.J.Pub.Health 42:1552.

Wheatley,G.M. (1955). New York State.J.Med. 55:3247.

General References.

Bell,E.T. (1950). Hypertension. Minnesota Press.

Best,C.H.,and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. The Williams and Wilkins Co.,Baltimore. 10th Ed.

Harrison,T.R. et al. (1954). Principles of Internal Medicine. The Blackiston Div.,McGraw-Hill Book Co. 2nd Ed.

Schroeder,H.A. (1953). Hypertensive Diseases. Lea and Febiger, Philadelphia.

Sunderman,F.W., and Boerner,F. (1950). Normal Values in Clinical Medicine. W.B.Saunders Co.,Philadelphia and London.

Wright,S. (1952). Applied Physiology. Oxford University Press. 9th Ed.

*Original article not read.

SCREENING FOR CARDIOVASCULAR DISEASE

Section C.
Chapter 3.

SCREENING FOR CARDIOVASCULAR DISEASE

This chapter will discuss the use of screening tests, other than measuring blood pressure, in detecting cardiovascular disease.

The Chest X-Ray

The chest x-ray has been the most widely used screening test for cardiac conditions. The miniature chest films, primarily taken for the early detection of pulmonary tuberculosis, have been read for cardiovascular conditions at little extra expense.

No specific shape of the heart shadow is pathognomonic of a particular kind of cardiac disease (Greening, 1954). The chest x-ray film will often provide only suggestive evidence of a heart abnormality. Greening (1954) emphasizes that a clinical examination and frequently an electrocardiogram are needed to interpret the roentgenographic findings.

Sosman (1947) and Ungerleider (1954) state that the chest x-ray can show the size of the heart better than any other method. However, cardiac enlargement can be present when the postero-anterior chest film appears normal. Moreover, slight scoliosis or slight rotation of the patient may result in a falsely positive reading of heart enlargement. Since heart disease may be present without cardiac enlargement, it is to be expected that the chest x-ray will be a screening method with low sensitivity and low specificity.

Special Readers for Heart Disease

Rutstein et al.(1951) used pairs of radiologists to look only for abnormalities of the heart and great vessels. These

special readers found almost three times as many cases of heart disease as did the regular readers, who worked individually and looked for all intra-thoracic conditions. Although the special readers increased the sensitivity of the test three-fold, the specificity of the test was unchanged.

More recently Anderson (1955) suggests that the yield of heart disease cases can probably be increased at least ninefold, using two readers paying particular attention to cardiovascular shadows. However, a joint committee (1953) of the American Public Health Association, National Tuberculosis Association and the U.S.Public Health Service has recommended that the use of special readers for heart disease does not appear justified.

Relative Accuracy of Miniature and Large Films

Rutstein et al.(1951) found that 14" by 17" films rarely provide information that cannot be obtained from the miniature films. Maclean and Rogen (1949) support this finding in analysing the results of the Mass Radiography Unit in Glasgow.

Comparison of Chest X-ray and Clinical Findings

Ungerleider (1954) states that "..physical examination is quite useless in detecting early enlargement of the heart." Recent findings of the Albany Cardiovascular Health Center confirm this statement. James and Hilleboe (1955) found that of 179 persons with chest x-rays suggestive of cardiac enlargement, 15 were found on physical examination to have this enlargement. Moreover, an additional 51 persons, whose heart shadows appeared normal in size, were believed by the clinicians to

have enlarged hearts. James and Hilleboe (1955) suggest that the routine physical examination, even when carried out by experienced clinicians, is of doubtful value in detecting cardiac enlargement.

Findings of Other Surveys

Rutstein et al.(1951) and Williamson et al.(1951) found that the yield of previously unknown cases of cardiovascular disease increased with age, and that there appeared to be a true increase in the precision of film reading with increase in age of the subject. The latter finding can be explained by the possibility that the type of cardiovascular disease detected may change with increase in age. For instance, Kurlander et al. show that the chest x-ray is most efficient in picking up arteriosclerotic heart disease, and least efficient in detecting hypertension without heart involvement.

Rutstein and his co-workers also found that the yield of the method was twice as high in females as in the same age group of males. The yields in non-whites were double those for whites in all age groups. Comstock (1953) found that this racial difference increased greatly with increasing age in a Georgia survey.

Kurlander et al. (1955), studying hospital outpatients, found that the chest x-ray detected 54 per cent of the previously known cases of heart disease and hypertension, and 34 per cent of the previously unknown cases. Weinerman et al. (1952), screening dock workers in San Francisco, were unim-

pressed by the use of the chest x-ray, which missed many cardiac cases and which only detected cases already screened out by the blood pressure or electrocardiogram tests.

The results of the Richmond, Virginia, Multitest Clinic were also unimpressive (American Medical Association, 1955). Of 2,148 positive screenees, 433 were known to have confirmed abnormalities, of which 55 were previously unknown.

Procedure in the Baltimore Clinic

A 70 mm. postero-anterior chest x-ray was carried out on most persons attending the clinic. A considerable number of technically unsatisfactory films were taken during the first weeks of the clinic. Such screenees were informed of this happening, and invited to return for a repeat miniature film if they desired to do so.

A chest physician read the films for the presence of tuberculosis, heart disease and other chest conditions. According to the wording of the physician's report, the reading was classified as normal, doubtful or abnormal for heart disease. "Possible heart enlargement," for instance, was classified as doubtful; "enlarged heart" was classified as abnormal. Most doubtfuls and all abnormals were referred to their family physicians for diagnosis.

For a few cases in which a full size film was recommended, appointments were made at nearby health department clinics for the 14" by 17" film before referral to the personal physician.

Results of the Baltimore Clinic

The results are shown in Table 1 of this chapter. Since

Table 1 - Results of 70 mm. Chest X-ray in Screening for Heart Disease, and of Follow-up with the Physician, by Age and Sex of Persons Screened.

Screening and follow-up results	Total 1/ persons screened	AGE IN YEARS, BOTH SEXES 2/ <30	30-44	45-64	65+	SEX 3/ All ages Male	Female
Number screened	1767	403	750	496	106	818	948
Normal	1585	393	717	401	67	729	855
Doubtful or abnormal	182	10	33	95	39	89	93
Heart condition previously unknown	155	9	31	83	28	76	79
Follow-up with physician completed	67	3	10	39	13	32	35
Heart disease confirmed	29	0	2	22	5	13	16
Previously unknown to physician	8	0	1	5	2	4	4
Rate per 1000 persons screened							
Number screened	1000	1000	1000	1000	1000	1000	1000
Normal	897	975	956	808	632	891	902
Doubtful or abnormal	103	25	44	192	368	109	98
Heart condition previously unknown	88	22	41	167	264	93	83
Follow-up with physician completed	38	7	13	79	123	39	37
Heart disease confirmed	16	0	3	44	47	16	17
Previously unknown to physician	5	0	1	10	19	5	4

1/ Includes unknown age and sex. 2/ Includes unknown sex.
3/ Includes unknown age.

no significant differences were found between the two sexes, the different age groups are not shown separately for each sex.

Age. The yield of the test increases with each succeeding age group. Under 45 years, the yield is very poor. A marked improvement in yield occurs in those 45 years and over.

Sex. There are no significant differences between the two sexes.

Completion of follow-up Procedure. Positive screenees of 45 years and over are more faithful in visiting the family physicians for the diagnostic examinations. However, in no age group does this proportion reach 50 per cent.

Confirmed cases. About one-third of the confirmed cases were previously unknown to the personal physician, according to the physician follow-up forms.

Race. Table 2 shows that the proportion of positive results is markedly greater among non-white persons. The fact that the average age of the non-white group is lower makes the difference more significant. The racial differences occur in both sexes.

Table 2 - Results of 70 mm. Chest X-ray in Screening for Heart Disease, by Race and Sex of Persons Screened.

Screening result	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	1767	1369	652	717	383	157	226
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Normal	897	919	906	930	817	822	814
Doubtful and abnormal	103	81	94	70	183	178	186

1/ Includes unknown race and sex.

Summary

The 70 mm. chest x-ray is a useful method of screening for previously unknown heart disease. The yield is mainly confined to those 45 years and over, and is greater in the non-white race. It is not known at present, however, to what extent these cases benefit by their early discovery.

The Electrocardiogram and Cardiovascular Disease

The electrocardiogram is able to show the presence of abnormal heart rhythms and of certain disease states of the myocardium. The most convenient type of machine to use in screening is the direct writing electrocardiograph.

The original criteria for the normal electrocardiogram were based in 1912 on Lewis' and Gilder's study of 53 young men. The significance of apparently abnormal patterns has been determined by autopsy of subjects who have died, usually from heart disease, shortly after the electrocardiogram was taken.

It is well established that normal records are frequently obtained from patients with heart disease. Furthermore, abnormal records may be obtained from persons without other evidence of heart disease. It is questioned whether electrocardiographic abnormalities found on routine examination of apparently healthy subjects should be given the same interpretation as when found on patients with symptoms of heart disease. For instance, Thomas (1944) recommends that "until the limits of normal variation in the human electrocardiogram have been much more thoroughly explored, the diagnosis of heart disease in young persons should seldom be based on electrocardiographic findings alone, in the absence of clinical manifestations."

Wilson and others (1947) support this opinion, and stress that electrocardiographic abnormalities are not diseases. "They have no important bearing upon the life expectancy of the patient, or the extent to which his mode of life should be altered when there is reasonable doubt as to the nature of the

factor or factors responsible for them in that particular case.. A diagnosis which, after careful investigation, is not supported by other evidence can seldom be made with confidence on the basis of electrocardiographic data alone.."

Effect of Position on the Electrocardiogram

The appearance of the QRS complex and of the T waves may be less reliable when the subject is in the upright position (Sigler, 1938; Ylvisaker and Kirkland, 1940). However, Dawber et al. (1952) studied the lead I tracings of 104 subjects and concluded that there was no essential difference in the reliability of the upright or recumbent positions.

The cardiologist reading the tracings will, of course, take into account the position of the subjects. However, the proportion of falsely positive readings may be reduced by taking the tracings in the recumbent position.

Electrocardiogram in the Elderly

Luisada (1954) emphasises that it is impossible to draw a sharp line between normal and abnormal tracings, particularly in the older age groups. "About one-fourth of apparently normal old persons present more severe electrocardiographic abnormalities indicating myocardial damage...These abnormalities may not be connected with clinical symptoms or signs.."

The normal electrocardiogram of the elderly subject may show T waves lower in amplitude, QRS waves slightly wider and occasionally slurred at or near the base line, and frequently shows left axis deviation (Sunderman and Boerner, 1950). The cardiologist reading the screening electrocardiograms should therefore know the age of each subject, to help in interpreting

the significance of the changes.

The Electrocardiogram, Hypertension, and Left Ventricular Changes

Fishberg (1954) mentions that the majority of patients with essential hypertension detected incidentally and without symptoms have normal electrocardiograms for their age and body type. When the hypertension has progressed sufficiently far to produce symptoms, the percentage of persons with ECG abnormalities increases.

Gubner and Ungerleider (1943) state that left axis deviation occurs in a large proportion of normal subjects, and is of no significance as a sign of left ventricular hypertrophy. They consider that this hypertrophy is present when the left axis deviation is accompanied by certain other changes which the authors specify in their article. They found that the electrocardiogram was more sensitive than the teleoroentgenogram for detecting left ventricular hypertrophy. At times, however, the one test could be relatively normal while the other showed definite evidence of the hypertrophy.

James and Hilleboe (1955) used the criteria of Gubner and Ungleider in comparing the findings of the first 1,494 persons attending the Albany Cardiovascular Health Center. They found that, while a person with an ECG suggesting enlargement is more likely to have x-ray evidence of this enlargement, the correlation between the two tests is not good. "It can be provisionally concluded that the two tests either show different types or stages of enlargement, or that the human

interpretation of the tests is sufficiently inexact to cause this poor correlation."

Autopsy studies have shown that cardiac hypertrophy rarely can be detected on the chest x-ray until it is accompanied by cardiac dilatation. The first alternative conclusion of James and Hilleboe, therefore, is probably the correct one, that the ECG shows evidence of left ventricular hypertrophy, while the chest x-ray shows the more advanced stage of left ventricular dilatation. Both tests are insensitive in screening for hypertension, with the electrocardiogram becoming positive usually at an earlier stage than the chest x-ray.

Consistency of Electrocardiograph Readings

Four cardiologists in the Albany Cardiovascular Health Center independently read the 12-lead electrocardiograms for signs of left ventricular hypertrophy (James and Hilleboe, 1955). The electrocardiograph interpretation varied from one reader to another. It may possibly also vary with the same reader at different times, although this was not tested in the Albany study. This subject needs further investigation.

Electrocardiogram and Obesity

Simonson and Keys (1932) found that with increasing body weight there is a pronounced tendency for the ECG pattern to show a more horizontal position of the heart. This finding might be wrongly interpreted if the presence of obesity were unknown.

A weight correction appears desirable for a more precise definition of normal electrocardiograph standards and for the

evaluation of an abnormality. This is particularly relevant in considering ECG records on the borderline between health and disease (Annotation, 1954).

When the screenee is obese, it seems advisable to note the presence of obesity to assist the electrocardiogram reader in interpreting the tracing.

Electrocardiogram in Screening

Ferguson and O'Connell (1926) studied 1,812 midshipmen, of whom all except three were considered free from heart disease. 52 per cent of their electrocardiographi tracings were perfectly normal, while an additional large number showed changes of no importance. Because of the number showing changes which would not usually be regarded as variations of the normal, the investigators concluded that the electrocardiogram was not a useful screening test.

Wood et al. (1941) studied 229 executives and 153 students selected because of suspected or known heart disease, and 705 unselected students. They concluded that "in persons under 30, electrocardiographic study does not demonstrate abnormalities of importance from the military standpoint which are not discoverable on physical examination."This does not, however, discourage its use in a screening program which involves no physical examination.

Viscidi and Geiger (1943) found that 50 per cent of 5000 unselected young working adults had tracings beyond the limits of normal. Included in the abnormalities, however, were slurring of the QRS complex (in 36 per cent) and prolonged QT interval

(in 17 per cent). It was concluded that, using the authors' criteria of abnormality, the electrocardiogram is an unreliable screening test which produces too many false positives.

Realising that the previously used criteria for abnormality had probably been largely variations of the normal, Hall et al. studied the records of 2000 members of the Royal Canadian Airforce Aircrew who had passed the standard medical examination. They classified 0.45 per cent of the tracings as "definitely abnormal", and 14.9 per cent as "doubtful abnormal."

Dawber et al. (1952) analysed the electrocardiograph tracings of 2000 volunteers in a cardiovascular disease study. Each 12-lead electrocardiogram was read separately by four physicians, and the majority reading was used to classify each tracing as normal, doubtful or abnormal. The test correctly classified 89 per cent of persons free of heart disease, but missed 53 per cent of those with cardiovascular disease. This result cannot be regarded as satisfactory in practice, since it will refer wrongly 11 per cent of persons without heart disease.

A similar study was recently carried out on 2,252 Los Angeles City employees (Phillips et al.,1953; Phillips,1955). After a complete physical examination, 162 persons were found to have heart disease and 80 had potential or possible heart disease. 13 per cent of the 2,010 persons without heart disease had abnormal 12- or 13-lead tracings, again confirming the high proportion of false positives. 65 per cent of the 242 persons with definite or possible heart disease had abnormal

electrocardiograms, confirming the low sensitivity of the test.

Matheyson (1954) took electrocardiograms during the periodic examination of employees. 19 per cent of those employees diagnosed clinically as normal had abnormal or doubtful tracings; 55 per cent of those with the clinical diagnosis of cardiovascular disease had abnormal or doubtful tracings. Matheyson believed that "the electrocardiogram is a useful adjunct to the periodic health examination." His results suggest, however, that the electrocardiogram without the physical examination will give too many false positive results.

Studying hospital inpatients, Witham and Jones (1956) took lead I and three chest leads on 126 cases of known heart disease and 92 persons without cardiovascular involvement. 83 per cent of the cardiacs were correctly classified and 11 per cent of the "normals" were wrongly classified as abnormal. Witham and Jones found that "minor" abnormalities were important in detecting the cases, and also mainly responsible for the false positives.

The Richmond, Virginia, multitest clinic carried out 12-lead electrocardiograms on 3,179 screenees. Of 445 persons screened as doubtful or positive, 99 were confirmed to have cardiovascular disease (American Medical Association, 1955). The previously unknown cases numbered 27.

The screening electrocardiograms of 3,984 dock workers in San Francisco produced 666 positives (Weinerman et al., 1952). 301 were eventually diagnosed to have cardiovascular disease, of whom 182 were previously unknown.

The Lead I Tracing

Because of the time taken to carry out 12 leads and the

need for undressing, Dawber et al. (1952) determined the value of using lead I alone as a screening procedure. Lead I was chosen because of the relative constancy of the complexes in normal individuals. The tracings were read by two readers, with 90 per cent agreement. The remaining 10 per cent were reread in order to agree upon a common interpretation. Their results were as follows:

	ECG Classification		
Clinical Diagnosis	Normal	Doubtful	Abnormal
No heart disease	80.3%(89.3%)	9.2%(7.3%)	10.5%(3.4%)
Possible heart disease	63.5%(64.1%)	12.6%(20.5%)	23.9%(15.4%)
Heart disease	50.9%(53.1%)	19.3%(18.4%)	30.2%(28.5%)

The figures in parentheses are the equivalent findings when the 12-lead tracing is used. The proportion of clinically normal persons wrongly classified by the ECG increased from 11 per cent to 20 per cent when the lead I tracing alone was used. There is relatively little change, however, in the effect of the test on those with heart disease. Dawber and his co-workers concluded that if the electrocardiogram be used as a screening test, lead I appears to be the method of choice.

Two recent studies have been less favourable to the use of lead I as a screening test. Kurlander et al. (1955) found that 23 per cent of hospital outpatients without hypertension or heart disease had positive lead I tracings. This proportion is probably higher, however, than would be obtained on screening apparently well persons from the general population. Of those outpatients newly discovered to have heart disease or

hypertension, 40 per cent had positive tracings. These results suggest that the lead I tracing is a screening test of low sensitivity and low specificity, and is not a good method of screening out individuals with cardiovascular disease.

Weintraub (1955) studied 200 definitely abnormal 12-lead electrocardiograms, carried out mainly on elderly males. 26 per cent were found to have a normal lead I tracing. Half of the subjects in this category had an old posterior wall or postero-lateral infarction. Weintraub concluded that, if electrocardiography is to be used in screening, multiple leads are much preferable to lead I alone.

The Baltimore Clinic

The evidence appeared fairly clear that, even with 12 leads the ECG would miss many cases of heart disease. However, it was felt that electrocardiography would detect a number of cases of heart disease which would otherwise remain undiscovered, despite the low sensitivity of the test.

It also appeared fairly clear that the lead I tracing alone would result in too many false positives. To improve this low specificity in the Baltimore clinic, the three bipolar and three unipolar limb leads were taken with screenees in the supine position. No undressing was thus required.

The tracings, bearing the age and blood pressure readings of the screenees, were classified by a cardiologist as normal, doubtful or abnormal. Persons with doubtful readings were invited to return for 12-lead ECGs, which were finally used to classify the result as normal or abnormal. The screenee was referred when the 12 or 6 lead tracing was abnormal, or when

the 6-lead tracing was doubtful and the screenee failed to return for the rescreening test.

The cardiologist used the revised recommendations of the Framingham Epidemiology Study as a guide in classifying the tracings (Kurlander et al.,1955). These criteria are given below:

Element	Abnormal Feature
P wave	1. Inverted 2. Height - 1.1 mm. or over 3. Width - 0.11 sec. or over. 4. Notched. 5. Biphasic. 6. Isoelectric.
P-R Interval	7. 0.20 sec. or over.
Q wave	8. Width - 0.04 sec. or over. 9. Depth - 1.5 mm. or over.
QRS interval	10. 0.11 sec. or over.
R wave, height	11. Under 1.5 mm. 12. 15 mm. and over.
S wave, height	13. Greater than R, unless low voltage.
S-T segment, deviation	14. 1 mm. or more deviation above or below base line.
T wave, amplitude	15. Inverted. 16. Isoelectric. 17. Biphasic. 18. 0.8 mm. or less.
Rate	19. 120 or over.
Rhythm	20. Auricular fibrillation or flutter. 21. Paroxysmal auricular tachycardia. 22. Paroxysmal ventricular tachycardia. 24. Heart block. 25. Multifocal or coupled rhythm.

The following are normal: Bradycardia of any degree, except in the presence of heart block (23). Ventricular premature systoles, unless multifocal or coupled rhythm (24). Wandering pacemaker. Occasional auricular and ventricular premature systoles.

Results of the Baltimore Clinic

The results are shown in Table 3 of this chapter. Since no significant differences were found in the same age groups of the two sexes, these age groups are not shown separately for each sex.

Table 3 - Results of Electrocardiogram in Screening for Heart Disease and of Follow-up with the Physician, by Age and Sex of Persons Screened.

Screening and follow-up results	Total 1/ persons screened	BY AGE IN YEARS Both sexes 2/ <30	30-44	45-64	65+	BY SEX 3/ All ages Male	Female
Number screened	2020	438	836	600	131	923	1095
Normal	1773	423	776	482	81	809	962
Doubtful or abnormal	247	15	60	118	50	114	133
Heart condition previously unknown	194	13	55	90	33	88	106
Follow-up with physician completed	128	7	27	66	26	54	74
Heart disease confirmed	43	0	2	30	12	21	22
Previously unknown to physician	16	0	1	10	4	9	7
	Rate per 1000 persons screened						
Number screened	1000	1000	1000	1000	1000	1000	1000
Normal	878	966	928	803	618	877	879
Doubtful or abnormal	122	34	72	197	382	123	121
Heart condition previously unknown	96	30	66	150	252	95	97
Follow-up with physician completed	63	16	32	110	199	59	68
Heart disease confirmed	21	0	2	50	92	23	20
Previously unknown to physician	8	0	1	15	31	10	6

1/ Includes unknown age and sex. 2/ Includes unknown sex.
3/ Includes unknown age.

Age. The proportion of positive rests and the yield of the test increased with each succeeding age group. Like the chest x-ray for heart disease, the yield was very poor with individuals under 45 years. Indeed, this yield is so low that the screening electrocardiogram appears suitable for use only with persons 45 years and over.

Completion of follow-up procedure. The proportion of previously unknown positive screenees who sought diagnostic examination rose with each increase in age group above 44 years. Only in the age group 30-44 years was this proportion under 50 per cent.

Confirmed cases. Just over one-third of the confirmed cases were previously unknown to the personal physician, according to the physician follow-up forms.

False positive:true positive ratio. When all the referrals are included, this ratio was 3½:1. When only those referrals who took the diagnostic examination are included, the ratio was just under 2:1.

Table 4 - Results of Electrocardiogram in Screening for Heart Disease, by race and sex of Persons Screened.

Screening result	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	2020	1569	735	834	432	177	255
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Normal	878	887	882	891	847	848	847
Doubtful and abnormal	122	113	118	109	153	152	153

1/ Includes unknown race and sex.

Race. Table 4 shows that the proportion of doubtful and abnormal results was greater among non-white persons, even although the average age of the non-white group attending was lower. The racial differences occurred in both sexes.

Rescreening results. Fifty-six individuals with a 6-lead ECG classified as "doubtful" were invited to return for a 12-lead electrocardiogram. Of 41 who did so, 31 were classified as normal by the repeat reading, and 10 as abnormal. The considerable reduction in the proportion of false positives appeared to make this rescreening procedure well worth while.

Comparison with Other "Heart" Screening Tests

The three screening procedures for "heart disease" resulted in the discovery of 33 confirmed cases per 1000 screenees. The electrocardiogram was positive in 21 of these cases, and was the only positive test in at least 8 of these cases. The ECG was more sensitive than the chest x-ray and the questionary, which were positive in 16 and 9 of the cases respectively per 1000 screenees.

Although the electrocardiogram produced the greatest yield of the three tests, this fact cannot be used as the main criterion for judging its value. The cases discovered by the electrocardiogram may or may not be less responsive to treatment than are the cases found by the other methods. Further study is needed to determine to what extent such cases benefit by early discovery.

Summary

The 6-lead electrocardiogram is a satisfactory screening

test for heart disease in persons 45 years and over. The ECG was more sensitive than the other "heart" tests used in the Baltimore clinic, although other studies have shown that its sensitivity is low. It is not known to what extent the cases discovered by the electrocardiogram benefit by their early detection.

t. Brodman et al. (1951) emphasise that it does n

used on outpatients attending a teaching clinic
rpreted by physicians, the index identified 5.6 d
eas (e.g. eye condition, hernia, varicose veins)
these areas being confirmed on diagnosis; 0.4 dia
se was missed per patient; 0.3 diagnostic area wa
d by the index and found not to be present on exa
amination.

last finding was probably the most significant f
of view of the questionary's value in screening.
agnostic areas per 100 patients were wrongly iden
found not to be present on examination. This erro
give less than thirty patients, but the number wou

The Questionary in Cardiovascular Disease

The Cornell Medical Index is probably the best known of all questionaries in use at present. As originally described, this index contains 195 questions corresponding closely to those usually asked in a detailed and comprehensive medical interview. Its completion takes 10-20 minutes.

The Cornell Medical Index draws attention to areas in which the patient may have disease manifestations. These areas must then be investigated to determine what specific disease is present. Brodman et al. (1951) emphasise that it does not "make automatic diagnoses."

When used on outpatients attending a teaching clinic and when interpreted by physicians, the index identified 5.6 diagnostic areas (e.g. eye condition, hernia, varicose veins) per patient, these areas being confirmed on diagnosis; 0.4 diagnostic area was missed per patient; 0.3 diagnostic area was identified by the index and found not to be present on diagnostic examination.

This last finding was probably the most significant from the point of view of the questionary's value in screening. Thirty diagnostic areas per 100 patients were wrongly identified and found not to be present on examination. This error would involve less than thirty patients, but the number would still remain seriously high. Since these findings were based on hospital outpatients, however, the proportion of false positives may be smaller when the questionary is used on apparently well persons.

Sodeman (1951) found that the Cornell Medical Index missed nothing of importance when its results were compared with the history taken by an internist. If the questionary had been used for screening, however, it would have caused the needless referral of 20 per cent of the individuals studied.

Derryberry, commenting on the results of the T.V.A. multiple screening survey, believed that the Cornell Medical Index was long and not well suited to screening apparently well persons. (American Medical Association, 1955).

Thirty-two questions, modified from the Cornell Medical Index, were used at first in the survey of dock workers in San Francisco (Weinerman et al.,1952). However, when 77 per cent of the men gave one or more affirmative answers, and 41 per cent gave three or more, its use for referral was abandoned.

Getting and Lombard (1952) gave particulars of the results of the first three Health Protection Clinics in Massachusetts. A Self-Screener was used, which was similar to the Cornell Medical Index. The figures given below are based on 1,399 referrals on whom reports were obtained from the family physician.

Diagnostic area	Positive on Screening.	Number of False Positives: 1 True Positive
Overweight	583	0.8
Hypertension	361	1.7
Female Genitals	345	2.5
Prostate	75	5.3
Underweight	79	5.5
Heart	274	6.2
Diabetes	174	6.9
Among the completely unsatisfactory were		
Arteriosclerosis	96	95.0

Getting and Lombard concluded: "Therefore it is our opinion that for the majority of diseases the self screener or clinical history is reasonably good, but for a minority of the diseases it is of little value." To reach this conclusion the authors appear to feel that the wrong referral of a false positive is not a serious matter. Some arbitrary decision has to be reached, however, on when a test becomes unsatisfactory. If it is decided that more than three false positives for every true positive should not be tolerated, only three diagnostic areas come within this limit. Among the diagnostic areas beyond this limit is that dealing with heart disease.

Kurlander et al. (1955) used a questionary on hospital outpatients, and found it a poor method of differentiating between persons with and without cardiovascular disease.

Phillips, Chapman and Goerke (1953; Phillips, 1955) are probably the investigators who have published the most helpful investigation of the use of the questionary in detecting heart disease. They asked 2,252 individuals selected at random from 20,199 Los Angeles City employees to answer seven questions. An answer which indicated that a symptom was present was considered positive.

162 of the 252 individuals were found on clinical examination to have heart disease. One-third of the cases had previous knowledge of the condition, and are excluded from the figures given on the following page.

Second and third examinations were carried out on these same persons 12-18 months apart. The questionary findings were comparable in all three examinations. The authors pointed out

the large number of false positives to most of the questions, suggested that the second, third and fourth questions were the most useful, but made no definite conclusions or recommendations.

Question	Positive answer by Cardiacs	Positive answer by Normals
1. Can you walk a reasonable distance outdoors without trouble ?	7.5%	1.3%
2. Do you ever have distress, pain or an uncomfortable feeling in the chest while walking on the street or up inclines or steps ?	33.3%	4.5%
3. While walking are you forced to stop in order to rest ?	20.0%	2.8%
4. Have you noticed increasing or undue shortness of breath on exertion ?	44.0%	17.0%
5. Is your sleep disturbed because of: Coughing spells ? Difficulty in breathing when lying flat in bed ? Asthma attacks ? Choking sensation in the chest ?	6%	Not available
6. Have you ever had palpitation of the heart ?	25.6%	10.8%
7. Are your ankles swollen at bedtime ?	8.6%	4.2%

The Baltimore Clinic

Questions 2 and 3 above were used in the Baltimore clinic questionary. It was decided to refer only those persons who gave a positive answer to both questions. This procedure would reduce the number of false positive results, as well as reducing the sensitivity of the test. The results are shown in Table 6 of this chapter.

Age. The proportion of screenees giving a positive answer to both questions tends to increase with age.

Sex. Women of all age groups consistently have a higher pro-

Table 6 - Results of Symptomatic Screening for Heart Disease and of Follow-up with the Physician, by Age and by Sex of Persons Screened

Screening and follow-up results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YEARS				
		2/ Total	<30	30-44	45-64	65+	2/ Total	<30	30-44	45-64	65+
Number screened	1898	876	165	374	281	54	1019	253	404	285	68
Negative	1502	721	133	333	213	40	779	222	323	189	37
One question positive	283	119	30	37	44	8	163	26	61	56	19
Both questions positive	113	36	2	4	24	6	77	5	20	40	12
Heart disease previously unknown	78	17	2	0	14	1	61	4	19	32	6
Follow-up with MD. completed	42	9	1	0	7	1	33	1	8	20	4
Heart disease confirmed	17	6	1	0	4	1	11	0	1	8	2
Previously unknown to MD.	5	0	0	0	0	0	5	0	1	3	1
	Rate per 1000 persons screened										
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Negative	791	823	806	890	758	741	764	877	799	663	544
One question positive	149	136	182	99	157	148	160	103	151	197	279
Both questions positive	60	41	12	11	85	111	76	20	50	140	177
Heart disease previously unknown	41	19	12	0	50	19	60	16	47	112	88
Follow-up with MD. completed	22	10	6	0	25	19	32	4	20	70	59
Heart disease confirmed	9	7	6	0	14	19	11	0	3	28	29
Previously unknown to MD.	3	0	0	0	0	0	5	0	3	11	15

1/ Includes sex and age not stated. 2/ Includes unknown age.

portion of positive answers to both questions than men of the same age groups.

Completion of follow-up procedure. In general, the proportion of positive screenees who sought diagnostic examination increases with age. 54 per cent of the previously unknown positive screenees visited their personal physician.

Confirmed cases. Just under one-third of the confirmed cases were previously unknown, according to the physician follow-up forms.

False positive:True positive ratio. When all the referrals are included, this ratio is 3.6:1. When only those referrals who had the diagnostic examination are included, the ratio is 1.4:1.

Table 7 - Results of Symptomatic Screening for Heart Disease, by Race and Sex of Persons Screened.

Screening results	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	1898	1485	705	780	396	162	234
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Negative	791	803	826	783	750	815	705
One question positive	149	141	133	149	177	148	197
Both questions positive	60	55	41	68	73	37	98

1/ Includes race and sex not stated.

Race. Non-white women gave more positive answers to both questions than did the white women. No significant difference occurred in the men.

Comparison with other "Heart" Screening Tests

As was to be expected, the questionary was the least sensitive of the three tests, being positive in only 9 of the 33 confirmed cases per 1000 screenees. The number of cases who would have been missed if the questionary had not been used is not known. The yield of previously unknown cases was 2.6 per 1000 screenees, all of the cases being women.

Summary.

The questionary for heart disease, as used intthe Baltimore clinic, was the least sensitive of the three "heart" tests. When an electrocardiogram and chest x-ray are used, it seems doubtful that the questionary significantly increases the yield of cases. In the absence of the other two tests, however, the questionary was sufficiently specific to be of value in detecting a small proportion of the cases.

,P.A.L. (1954). Canad.Med.Ass.J. 70:632.

,B. (1955). California Med. 82:

,E.,Chapman, E.A. and Goerke,L.S. (1953). Amer.He 45:219.

REFERENCES

American Medical Association, (1955). A Study of Multiple Screening. Council on Medical Service.

Anderson,R.J. (1955). J.Chronic Dis. 2:427.

Annotation (1954). Nutr.Rev. 12:305.

Brodman,K. et al. (1951). J.Amer.Med.Ass. 145:152.

Comstock,G.W. (1953). New England J.Med. 248:1045.

Dawber,T.R.,Kannel,W.B.,Love,D.E., and Streeper,R.B. (1952) Circulation 5:559.

Ferguson,D.,and O'Connell,J.T. (1926). U.S.Nav.Bull.24:860.*

Fishberg,A.M. (1954). Hypertension and Nephritis. Lea and Febiger, Philadelphia.

Getting,V.A.,and Lombard,H.L. (1952). New York State Med.J. 52:2605.

Greening,R.R. (1954). Med.Clin.N.Amer. 38:1575.

Gubner,R.,and Ungerleider,H.E. (1943). Arch.Int.Med. 72:196.

James,G.,and Hilleboe,H.E. (1955). Amer.J.Public Health 45:140.

Joint Committee (1953). Public Health Rep. 68:415.

Kurlander,A.B.,Hill,E.H., and Enterline,P.E. (1955). J.Chronic Disease 2:427.

Luisada,A.A. (1954). Geriatrics 9:381.

Maclean,A.,and Rogen,A. (1949). British Med.J. 1:264.

Matheyson,F.A.L. (1954). Canad.Med.Ass.J. 70:632.

Phillips,E. (1955). California Med. 82:118.

Phillips,E.,Chapman,J.M., and Goerke,L.S. (1953). Amer.Heart.J. 45:319.

Rutstein,D.D.,Williamson,C.R.,and Moore,F.E. (1951). Circulation 4:641.

Sigler,L.H. (1938). Amer.Heart J. 15:146.

Simonson,E.,and Keys,A. (1932). Circulation 6:749.

*Original article not read.

Sodeman,W.A. (1951). Canad.J.Public Health 42:43.

Sosman,M.C. (1947). Roentgenographic Diagnosis of Cardiac Disease. Nelson New-Loose-Leaf Medicine. Vol.IV. P 581. Thomas Nelson and Sons.

Sunderman,F.W., and Boerner,F. (1950). Normal Values in Clinical Medicine. 1st Ed. W.B.Saunders Co.,Philadelphia and London.

Thomas,C.B. (1944). Bull.Johns Hopkins Hosp. 74:229.

Ungerleider,H.E. (1954). Indust.Med.& Surg. 23:525.

Viscidi,P.C., and Geiger,A.S. (1943). Amer.Heart J. 26:763.*

Weinerman,E.R. et al. (1952). Amer.J.Public Health 42:1552.

Weintraub,H.J. (1955). J.Amer.Med.Ass. 158:178.

Williamson,C.R.,Moore,F.E., and Rutstein,D.D. (1951). Circulation 4:652.

Wilson,F.N. et al. (1947). in Advances in Internal Medicine. Editors, Dock,W., and Snapper,I. Vol.2. P 1. Interscience Publishers, Inc.

Witham,A.C.,and Jones,H.B. (1956). American Heart J. 51:186.

Wood,F.C. et al. (1941). War Medicine 1:696.*

Ylvisaker,L.S.,and Kirkland,H.B. (1940). American Heart J. 20:592.

General References.

Best,C.H.,and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. The Williams and Wilkins Co. 6th Ed.

Dimond,E.G. (1954). Electrocardiography. The C.V.Mosby Co.

Dressler,W.,and Roesler,H. (1949). An Atlas of Electrocardiography. Charles C.Thomas.

Goldberger,E. Heart Disease. Lea and Febiger, Philadelphia. 2nd Ed. 1955.

Luisada,A.A. (1954). Heart. The Williams and Wilkins Co. 2nd Ed.

Massey,F.C. (1953). Clinical Cardiology. The Williams and Wilkins Co.

*Original article not read.

Schroeder,H.A. (1953). Hypertensive Diseases. Lea and Febiger, Philadelphia.

White,P.D. (1951). Heart Disease. The MacMillan Co.,New York. 4th Ed.

Wright,S. (1952). Applied Physiology. Oxford University Press. 9th Ed.

Section 6
Chapter 4

...cia is a deficiency in the number of red cells, ...
...globin or the relative volume of packed cells per ...
... (Wintrobe, 1954). A number of factors must be ...
...ing for anaemia.

...emoglobin fluctuations in the normal person ...
...during ... even under ordinary ...
...ion, but concentrations of red cells may vary in dif...
...s of the circulation. The proportion of red cells ...
...ibution is ... as a rule than to the heart and ...
...cells, by ... per cent."

Section C
Chapter 4

...ctuations of 1-2 gm. per cent are stated to occur ...
... short period of time (Dunn, 1931; Dawber, 1945), ...
...ly due to the release into the circulation of ... ...
...cells stored in the splenic pulp. Exercise and emo... ...
... such as fear and excitement produce haemoconcentra... ...
...ay produce an increase in plasma volume, with a ...
...globin concentration.

... a definite and the postprandial fall, diurnal var... ...
... and seasonal changes (Wintrobe, 1951). Except when ...
...es in altitude exist, there is no convincing evi... ...
... that geographic factors affect the haemoglobin le...
..., 1951).

SCREENING FOR ANAEMIA

SCREENING FOR ANAEMIA

Anaemia is a deficiency in the number of red cells, amount of haemoglobin or the relative volume of packed cells per unit of volume (Wintrobe, 1954). A number of factors must be considered in screening for anaemia.

Haemoglobin Fluctuations in the Normal Person

Banting and Best emphasise that, "even under ordinary conditions, the concentration of red cells may vary in different parts of the circulation. The proportion of red cells in the capillaries is greater as a rule than in the heart and larger vessels, by around 12 per cent.."

Fluctuations of 1-2 gm. per cent are stated to occur within a short period of time (Doan, 1931; Fowler, 1949), apparently due to the release into the circulation of a supply of red cells stored in the splenic pulp. Exercise and emotional factors such as fear and excitement produce haemoconcentration. Recumbency produces an increase in plasma volume, with a fall in haemglobin concentration.

Less definite are the postprandial fall, diurnal variations and seasonal changes (Wintrobe, 1951). Except where differences in altitude exist, there is no convincing evidence to show that geographic factors affect the haemoglobin level (Wintrobe, 1951).

In screening, the fluctuations would seem to produce errors in the direction of false negative results -- the mildly anaemic persons being classified as "normal" because of a temporary rise in haemoglobin concentration. The factors men-

tioned so far should not require an alteration in the screening level.

Sex Difference in Haemoglobin Concentration

Widdowson and McCance (1936) gave 100 mg. iron daily for 2-3 weeks to 15 men and 16 women. There was a sharp rise in the haemoglobin concentration of the women to reach approximately the accepted "normal" level for men. The men showed only a slight rise in the haemoglobin concentration. The authors believed that there was no true sex difference in haemoglobin levels, and that the frequently found lower levels in women should be regarded as "mildly pathological."

Yudkin (1944) studied the blood of recruits to the Women's Auxiliary Air Force in Great Britain. After six months or more on a diet of high iron content the haemoglobin concentrations of the women had risen to reach a value acceptable as "normal" for men.

Twenty-two apparently normal women of child bearing age took 300-346 mg. ferrous iron daily for eight weeks. Vahlquist (1950) found that the haemoglobin concentration did not increase beyond the values in another comparable group of forty women, and remained markedly less than the haemoglobin concentration of men.

Garry et al. (1954) made two series of observations on the concentration of haemoglobin in the blood of men and of women students attending the University of Glasgow. One year's administration of iron supplement tablets failed to diminish the difference in haemoglobin concentrations in men and women. The authors concluded that "..the difference between the sexes

cannot be attributed to an inadequate intake of iron on the part of the women."

These selected references show that there is no uniform opinion concerning the different haemoglobin concentration in each sex. However, most family physicians use lower limits of normal for women than for men. It would therefore appear wise to use a lower screening level for women.

Haemoglobin Concentrations in Pregnancy

Widdowson (1939) found that large doses of iron halted the fall in haemoglobin concentration which usually occurs in pregnancy. Benstead and Theobald (1952), Davis and Jennison (1954), and Gatenby and Lillie (1955) had similar findings and concluded that the so-called "physiological anaemia" of pregnancy was due to the inadequate intake of iron.

Fisher and Biggs (1955) studied 104 patients who took iron regularly from the fourth month of pregnancy. Twelve cases showed a general downward trend in haemoglobin levels during pregnancy, with a sharp post-partum rise. The remaining 92 patients had a haemoglobin level around 14.5 gm. per cent at 38 weeks and remained at much the same level for six or more weeks after delivery. The authors concluded that the frequently reported fall in haemoglobin during pregnancy can usually be prevented by administration of iron.

Holly (1955) supports this finding. He feels that 12 gm per cent should be regarded as the lower limit of normal in pregnancy, since 80 per cent of mothers remain above this level when given supplemental iron.

Garry et al. (1954) feel that such studies are not convincing, and state: "During pregnancy the concentration of

haemoglobin in the blood of the mother falls still further. There seems to be an undoubted increase in the total blood volume, but the increase in the volume of the erythrocytes is not so great as the increase in the volume of the plasma. The resulting decrease in the concentration of the heamoglobin is regarded not as an anaemia but as an 'oligocythemic hypervolemia'.."

However, and Editorial (1955) comment on the study by Fisher and Biggs (1955) states: "The attainment and maintenance of a high haemoglobin level during and after pregnancy not only increases the margin of safety for mother and foetus, but it also ensures the woman a quicker recovery from childbirth and gives her more energy to resume her household duties and to enjoy her motherhood." It therefore seems reasonable to use one screening level for all women, whether pregnant or not. While this procedure will not meet with the unanimous approval of the family physicians, it is administratively simple and has considerable medical evidence for its support.

Haemoglobin Concentrations in the Elderly

Several studies suggest that the haemoglobin levels remain fairly constant in women after the menopause, but that the levels of men over sixty years tend to fall and frequently reach the female levels in the seventies (Hawkins et al.,1954; Hawkins, 1956). There is no general agreement that this fall is a "normal" occurrence, since an inadequate dietary intake may be involved. However, there is also no agreement that elderly men with these lower levels should be given or will

respond to anti-anaemic therapy. Until the haemoglobin fall in elderly males has been studied more thoroughly, it seems reasonable to use the same screening level for males of all ages.

The Limits of Normal

Probably the most widely used figures for the normal range of haemoglobin concentration are those suggested by Wintrobe (1951). Wintrobe gives the normal range for males as 14.0-18.0 gm. per cent, with a mean of 16 gm. per cent; that for females as 12.0-16.0 gm. per cent, with a mean of 14. gm. per cent.

Most textbooks of clinical pathology recommend values in this region, with two different ranges usually given for the two sexes. Sometimes the range quoted is considerably wider, such as 13-20 gm. per cent for males and 11-18 gm. per cent for females (Levinson and MacFate, 1951). Occasionally a text-book, such as that by Todd, Sanford and Wells (1953), suggests that the same range should be used for both sexes -- in this case 13.5-17.5 gm. per cent.

Hawkins et al. (1954) studied the haemoglobin values of 1,308 males and 1,424 females in Halifax, Nova Scotia. The values in women remained fairly constant above age 30 years. The male values remained fairly constant between the ages of 18 and 50 years. Beyond this age, the mean values showed a steady fall. Hawkins et al. mention that not more than 5 per cent of males or females in any age group had values below 12 gm. per cent and 10.0 gm. per cent respectively. These

levels would appear reasonable for use in screening for anaemia.

D'Alonzo and Rodgers (1955) studied the haemoglobin values of 17,956 adult men. 7,679 had values 14.0 gm. per cent or below, and 174 were at or below 12.0 gm. per cent. A total of 103 of these men were ultimately diagnosed to have anaemia or another condition accompanied by anaemia. It would therefore appear that the frequently accepted lower limit of normal of 14.0 gm. per cent for the haemoglobin level in men will screen off large numbers of persons who will not be diagnosed as having anaemia.

Racial Difference in Haemoglobin Concentrations

Lower haemoglobin levels are more frequently found in negroes than in white persons in the United States. This difference is usually attributed to a greater prevalence of anaemia in negroes, and not to a true racial difference.

Youmans et al. (1950) surveyed 81 per cent of persons aged 16 years and over living in Wilson County, Tennessee. They apparently used 13 gm. per cent and 11 gm. per cent as the lower limits of normal for males and females respectively. They found that anaemia was present in 14 per cent of the population, with the percentage being considerably greater in negro and considerably less in white persons. Their results, and those of Milam and Muench working in North Carolina (1946) illustrate the fact that the use of 14 and 12 gm. per cent as the levels for referral of males and females respectively will result in a considerable proportion of the population, especially the negro population, being placed in the abnormal classification.

However, although the use of the same screening levels for coloured and white persons will result in the referral of a greater proportion of negroes, it is possible that this referral will be beneficial.

Value of Early Detection

Although the screening levels suggested here are lower than those usually used as the lower limits of normal, they are still considerably higher than the levels frequently found when anaemia is first diagnosed in the physician's office.

Whitby and Britton (1950) state that the haemoglobin level is between 6 and 9 gm. per cent when advice is first sought. Wintrobe (1951) suggests 6 to 10 gm. per cent, while Witts (1950) mentions that the haemoglobin concentration may fall below 6 gm. per cent before discovery of the anaemia. In the anaemia of pregnancy, Gatenby and Lillie (1955) found that all patients were clinically anaemic only when the haemoglobin was below 5.9 gm. per cent.

Even when comparatively low screening levels are used, the anaemia is likely to be detected considerably before symptoms occur to cause the individual to seek medical advice. It is also likely that the individual will benefit by the earlier discovery and treatment of the anaemia.

Methods of Haemoglobin Determination

The Tallqvist method uses absorbent paper on which a drop of blood is placed. The colour produced is compared with a carefully printed colour scale. This method has too great an

error for screening, this error being stated to be 20-40 per cent.

Previous screening programs have used the Dare haemoglobinometer. This method uses undiluted blood, which is matched directly with a colour scale. The instrument is expensive, and the errors may be as great as 20 per cent (Todd, Sanford and Wells, 1953). This method does not appear satisfactory for use in screening.

The Haldane Gowers method converts the haemoglobin to carboxyhaemoglobin. Since this procedure needs a source of carbon monoxide to be bubbled through the haemoglobin solution, it is somewhat complicated for screening large numbers of persons. The average technician is liable to an error of 10 per cent, although indifferent users produce errors of 20 per cent or more.

Several acid haematin methods, such as those of Sahli or Haden-Hauser, use blood which has been allowed to stand for a certain length of time in decinormal hydrochloric acid. The colour produced is then compared with a standard. The rate of acid haematin formation is stated to vary (Dyke, 1951), and requires about 24 hours for completion. However, 95 per cent of the colour is attained after ten minutes, and the change is insignificant after half an hour (Wintrobe, 1954). The reading has therefore to be carried out at a specific time after the test is begun, and a delay of some minutes will cause the colour to change. The usual error is $\pm$ 10 per cent (Miller, 1955; Cartwright, 1954), although it rises to $\pm$ 30 per cent when the technique and equipment are not standardised. The

acid haematin tests do not appear to be very satisfactory for use in screening.

A number of photoelectric colourimetric methods have been developed, which are fairly rapid and have a range of error of about ± 5 per cent. These methods would appear to be useful in screening, especially when the actual haemoglobin values are desired. A test with this degree of accuracy will permit the use of screening levels which are closer to the "lower limits of normal" without referring too many false positives.

Copper Sulphate Specific Gravity Method

Transfusion services use this test in many parts of the world, and believe it satisfactory for the rapid examination of large numbers of individuals. The test measures the specific gravity of blood, which is related to the haemoglobin content.

The specific gravity of plasma varies around 1.027, depending on the plasma protein concentration. The red cells have a specific gravity of around 1.097 (Houssay et al.,1955). Since haemoglobin makes up from 80 to 90 per cent of the total solids of the red cell, the specific gravity of whole blood will be mainly affected by the haemoglobin concentration, although also influenced to a lesser extent by other constituents.

When the specific gravity of the blood plasma is determined separately, the haemoglobin estimation is stated to have an error of ± 2 per cent. When the plasma specific gravity is not estimated separately, it is assumed to be about 1.027 and the resultant haemoglobin estimation is stated to have an error of ± 10 per cent. This simplified method is the usual screening procedure.

The test is carried out by allowing a small drop of whole blood to fall into copper sulphate solutions of known specific gravity. On entering the solution each drop becomes encased in a sack of copper-proteinate, and remains as a discrete drop without change of specific gravity for 15 or 20 seconds.

Within five seconds the momentum of the fall is lost and the drop then either begins to rise, or becomes stationary, or continues to fall. A drop lighter than the test solution will rise, perhaps only a few millimetres, and may begin to sink immediately afterwards.

If the drop is of the same specific gravity as the test solution it will become stationary for a few seconds and then fall. If the drop is heavier than the solution it will continue to fall on entering the solution. The behaviour of the drop in the ten seconds after it has lost its momentum of fall indicates whether the drop is lighter or heavier than the test solution.

When the actual plasma specific gravity is within normal range, the error caused by taking the fixed value of 1.027 is $\pm$ 4 per cent. When a pathological condition has caused the plasma specific gravity to go beyond the normal range, the haemoglobin estimation error is $\pm$ 9 per cent. These figures give the error when the test is carried out with complete accuracy. An additional figure must be added for the error involved in carrying out the test.

In the Baltimore clinic, three different solutions were used with specific gravities corresponding to haemoglobin concentrations of 12.3,11.3 and 10.0 gm. per cent. When a drop

of blood remained stationary or rose in the 11.3 gm. per cent solution, the test was regarded as positive for men. The 10.0 gm. per cent solution was used for the referral of women. The 12.3 gm. per cent solution, which is the value used to determine elegibility for donating blood in the United States, was included only for research purposes.

The Control Division of the Bureau of Laboratories of the Maryland State Department of Health retested 116 of the venous blood specimens. An oxyhaemoglobin photoelectric colourimetric procedure, which was likely to give results within 5 per cent of the correct value, was used. A comparison of the results obtained by the two methods suggests the following:

1. The actual screening levels used were about 1 gm. per cent higher than had been realised. When a specimen of blood with a haemoglobin concentration of 11.3 gm. per cent was dropped into the 11.3 gm. per cent solution, it would theoretically have an equal chance of being classified as above or below that level, since no other result could be recorded. However, because a drop which remained stationary for a few seconds was arbitrarily classified as "below" the level, the specimen actually had to be 1 gm. per cent greater than the solution before there was a 50 per cent chance of being classified as "above". In practice, therefore, the women had been screened at 11.0 gm. per cent and the men at 12.3 gm. per cent. Bowdoin (1951) has published a similar finding.

2. As long as the blood specimen was within 3 gm. per cent of the actual screening level, there was a chance that the specimen would be wrongly classified. For instance, 2 per cent

of specimens with a haemoglobin concentration of 9.3 gm. per cent, when dropped into the 11.3 gm. per cent solution, would be wrongly classified as "above" the screening level.

The fact that the technicians in Baltimore had no previous experience with this test may have resulted in the error being larger than usual. It is evident, however, that the error is a very considerable one. The results support Liddelow's (1953) conclusions that this test is unreliable for accurate work and of value only in obtaining an approximate estimate in an urgent case.

In summary, the copper sulphate specific gravity method of determining haemoglobin concentration is simple, rapid and inexpensive. Its large range of error, however, suggests that the test should be replaced in future screening programs by a photoelectric colourimetric method. The specific gravity method will refer a considerable proportion of false positives unless the screening levels used are well below the "limits of normal" for haemoglobin concentration.

Packed Cell Volume in Screening.

Clinicians widely use a haematocrit test for packed cell volume to determine if a more thorough blood examination is required. The packed cell volume is low when the number of cells is reduced, the amount of haemoglobin is diminished, or when both changes are present. Either of these findings indicate the need for further examination.

Since duplicate samples centrifuged at the same time will

agree within ± 1 per cent the test is as accurate as the most exact of the haemoglobin tests. Its main drawbacks in screening are the amount of blood required and the need to have several centrifuges working continuously.

A method is now available, however, which uses only a small amount of blood centrifuged for five minutes at high speed in a capillary tube which is discarded at the end of the test. This method is stated to have an error of ± 2 to 4 per cent (Miller, 1955), and would appear to be quite practicable for use in screening.

Previous Screening Programs for Anaemia

The Richmond, Virginia, Multitest Clinic (American Medical Association, 1955) used the Dare haemoglobinometer to determine haemoglobin concentration. Both sexes were referred if below 12 gm.per cent. Of the 4,967 positive screenees, 1,034 were confirmed to have anaemia by their personal physicians. Half of the confirmed cases were previously unknown.

The Indianapolis, Indiana, multiple screening program (Carroll, Kurlander and Nester, 1954) used a photoelectric colourimetric method, with 12.5 and 11.0 gm. per cent as the screening levels for males and females respectively. 923 positive screenees produced 220 confirmed cases, 96 of whom were previously unknown.

The dock workers' survey in San Francisco, California, produced only five positive screenees below 12.3 gm. per cent, with one newly discovered confirmed case (Weinerman et al., 1952).

Derryberry (1955) comments on the T.V.A. multiple screening survey in Alabama that the inclusion of a haemoglobin estimation in the screening procedure did not appear to be worth while. Since 1.4 per cent of the screenees were diagnosed as having anaemia or other blood conditions, however, others may disagree with his conclusion.

Collen and Linden (1955) found that 86 of 1000 persons screened below 13 gm. per cent by the copper sulphate specific gravity method. Of 39 who were retested by the Klett colourimetric procedure, 39 per cent were confirmed to be below this level.

The Baltimore Clinic

Men were referred if the drop of blood remained stationary or rose in the solution corresponding to a haemoglobin concentration of 11.3 gm. per cent. The 10.0 gm. per cent solution was used to refer women. This procedure actually resulted in the referral of men below 12.3 gm. per cent and of women below 11.0 gm. per cent. These latter values will now be referred to as the actual screening levels.

The results are shown in Table 1 of this chapter.

Age. The proportion of positive tests decreased with age in women.

Sex. The majority of positive tests and all the confirmed cases occurred in women. A screening level of 12.3 gm. per cent does not produce a satisfactory yield of cases in men.

Completion of follow-up procedure. Less than 50 per cent of the positive screenees had a diagnostic examination.

Table 1 - Results of Screening for Anaemia and of Follow-up with the physician, by Age and Sex of Persons Screened.

Screening and follow-up results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YEARS				
		2/ Total	<30	30-44	45-64	65+	2/ Total	<30	30-44	45-64	65+
Number screened	1980	914	171	402	283	56	1063	260	421	300	71
Normal	1948	911	171	400	282	56	1034	247	408	293	71
Abnormal	32	3	0	2	1	0	29	13	13	2	0
Anaemia previously unknown	29	3	0	2	1	0	26	11	12	2	0
Follow-up with MD. completed	13	0	0	0	0	0	13	6	5	2	0
Anaemia confirmed	8	0	0	0	0	0	8	4	3	1	0
Previously unknown to MD.	7	0	0	0	0	0	7	4	2	1	0
	Rate per 1000 persons screened										
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Normal	984	997	1000	995	996	1000	973	950	969	993	1000
Abnormal	16	3	0	5	4	0	27	50	31	7	0
Anaemia previously unknown	15	3	0	5	4	0	25	42	29	7	0
Follow-up with MD completed	7	0	0	0	0	0	12	23	12	7	0
Anaemia confirmed	4	0	0	0	0	0	8	15	7	3	0
Previously unknown to MD.	4	0	0	0	0	0	7	15	5	3	0

1/ Includes unknown age and sex 2/ Includes unknown age.

False positive:true positive ratio. This is 2½:1 when all the referrals are included, and 3:4 when only those referrals are included who completed the follow-up examination.

Previously unknown cases. Almost all of the confirmed cases were previously unknown to the physician.

Race. Table 2 shows that the proportion of positive tests is greater among non-white persons than among white persons of the same sex.

Table 2 - Results of Screening for Anaemia, by Race and Sex of Persons Screened.

Screening results	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	1980	1539	727	812	422	176	246
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Normal	934	986	997	977	974	994	959
Abnormal	16	14	3	23	26	6	41

1/ Includes unknown race and sex.

Summary

The screening of men for anaemia at a level of 12.3 gm. per cent is a non-productive test. The screening of women at 11.0 gm. per cent gives a reasonable yield of previously unknown cases. It is suggested that future screening programs should replace the copper sulphate specific gravity method with a more accurate photoelectric colourimetric method.

REFERENCES

American Medical Association (1955). A Study of Multiple Screening. Council on Medical Service.

Benstead,N., and Theobald,G.W. (1952). British Med.J. 1:407.

Carroll,B.E.,Kurlander,A.B., and Nester,H.G. (1954). Public Health Reports 69:1180.

Cartwright,G.E. (1954). Diagnostic Laboratory Hematology. Grune and Stratton, Inc. 1st Ed.

Collen,M.F., and Linden,C. (1955). J.Chronic Dis. 2:400.

D'Alonzo,C.A., and Rogers,S.M. (1955). Indust.Med.& Surg. 24:75.

Davis,L.R.,and Jennison,R.F. (1954). J.Obstet.Gynaec.Brit.Emp. 61:103.

Derryberry,O.M., in A Study of Multiple Screening. Council on Medical Service, American Medical Association. P 33. 1955.

Doan,C.A. (1931). Medicine 10:323.*

Dyke,S.C. (1951). Recent Advances in Clinical Pathology. The Blackiston Co. 2nd Ed.

Editorial (1955). British Med.J. 1:405.

Fisher,M.,and Biggs,R. (1955). British Med.J. 1:385.

Fowler,W.M. (1949). Hematology. Paul B. Hoeber, Inc. 2nd Ed.*

Garry,R.C., et al. (1954). British J. Nutr. 8:253.

Gatenby,P.B.B.,and Lillie,E.W. (1955). Lancet 1:740.

Hawkins,W.W. (1956). J.Amer.Ger.Soc. 4:24.

Hawkins,W.W.,Speck,E., and Leonard,V.G. (1954). Blood 9:999.

Holly,R.G. (1955). Obstet.& Gynec. 5:563.

Houssay,B.A., et al. (1955). Human Physiology. McGraw-Hill Book Co.,Inc. 2nd Ed.

*Not read in the original.

Levinson,A.L.,and MacFate,R.P. (1951). Clinical Laboratory Diagnosis. Lea and Febiger. 4th Ed.

Liddelow,B. (1953). Med.J.Australia 2:565.

Milam,D.F.,and Muench,H. (1946). J.Lab.Clin.Med. 31:878.

Miller,S.E. (1955). A Textbook of Clinical Pathology. The Williams and Wilkins Co. 5th Ed.

Todd,J.C.,Sanford,A.H.,and Wells,B.B. (1953). Clinical Diagnosis by Laboratory Methods. The W.B.Saunders Co. 12th Ed.

Vahlquist,S. (1950). Blood 5:874.

Weinerman,E.R.,et al. (1952). Amer.J.Public Health 42:1552.

Whitby,L.E.H.,and Britton,C.J.C. (1950). Disorders of the Blood. The Blackiston Co.

Widdowson,E.M. (1939). Lancet 260:131.

Widdowson,E.M.,and McCance,R.A. (1936). J.Hyg.Camb. 36:13.*

Wintrobe,M.M. (1951). Clinical Hematology. Lea and Febiger.

Wintrobe,M.M. (1954) in Practice of Medicine. F.Tice, Editor. Vol.VI. W.F.Prior Co.

Witts,L.J. (1950). in Textbook of the Practice of Medicine. F.W.Price, Editor. Oxford University Press.

Youmans,J.B.,et al. (1950). Amer.J.Medical Science 219:30.

Yudkin,S. (1944). British Med.J. 2:403.

General References.

Albritton,E.C. (1952). Standard Values in Blood. W.B.Saunders Co.

Best,C.H.,and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. The Williams and Wilkins Co. 6th Ed.

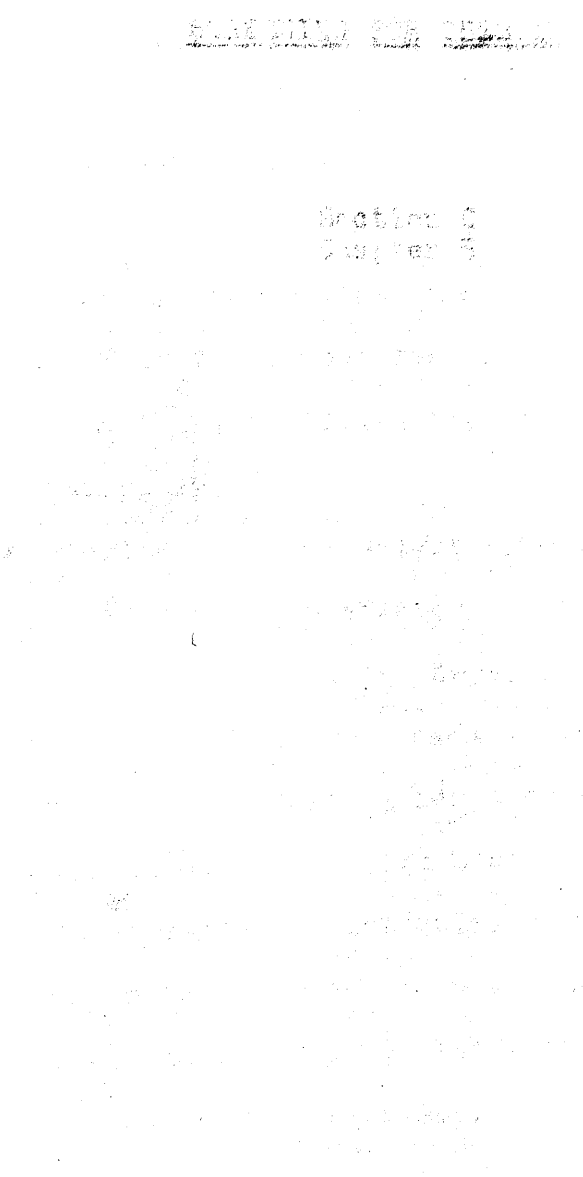
Goodale,R.H. (1954). Clinical Interpretation of Laboratory Tests. F.A.Davis,Co.

Sunderman,F.W.,and Boerner,F. (1950). Normal Values in Clinical Medicine. W.B.Saunders Co.

*Original article not read.

Simmons,J.S.,and Gentzkow,C.J. (1955). Medical and Public Health Laboratory Methods. Lea and Febiger.

Wright,S. (1952). Applied Physiology. Oxford University Press. 9th Ed.



SCREENING FOR OBESITY

Section C
Chapter 5

SCREENING FOR OBESITY

Obesity is present in a person who has an excessive amount of adipose tissue. While the obese person is usually overweight, not all overweight persons are obese. An individual may be overweight, for instance, when his musculo-skeletal system is unusually well developed. Such a person is not necessarily obese.

Overweight is closely associated with excessive mortality from several chronic diseases. Persons more than 25 per cent above "normal" weight have more than twelve times the mortality rate from diabetes; three times that from cerebral haemorrhage and apoplexy; four times that from nephritis; and two times that from organic heart disease than do slightly underweight persons (Breslow, 1951).

The comparison of body weight with standard age-height-weight tables is not a satisfactory method of determining the presence of obesity (Brozek and Keys, 1950). At present, however, no other method has been made sufficiently standard or simple to be used in screening for obesity.

Skin fold measurements have been extensively used in recent nutrition surveys, particularly in Canada. The calipers for measuring the skin folds, however, have not yet been standardised for this procedure (Edwards et al.,1955). Moreover, there is no general agreement on the sites to be measured or on the range of normal values.

Hammond (1955) has advocated expressing fat measurements relative to the x-ray thickness of uncompressed fat. This method would be useful in screening if the chest x-ray could

be used for these measurements. Berry and Nash (1955) have shown that there is a correlation between chest x-ray measurements and caliper measurements, but that the x-ray can have a considerable margin of error in individual cases.

Available Height-Weight Tables

The Metropolitan Life Insurance Company (1943) tables have been widely used in the United States. These tables give one set of values for adult men, and one for adult women. The values are based on the belief that, while the body weight frequently increases with age, this increase is not biologically desirable. Indeed, there is now evidence to suggest that the lean body mass decreases gradually with age, and that the body fat must increase with age to make the body weight remain constant (Murphy, 1955).

The Metropolitan Life tables, developed in 1912, used the measurements taken on some 200,000 males and on a smaller number of females. (There is evidence that many of the weights were estimates only, and were not actually obtained by weighing the individuals.) The distribution of body weights was typically skewed towards the higher values, so that the mean was greater than the median. This type of distribution has been confirmed in other studies.

The original figures underwent a number of manipulations in the development of the tables (Keys and Brozek, 1953). The figures for men of 30 years and women of 25-29 years were used. Average weights were calculated for each inch of height, and

then somewhat reduced to approach the median values. The weights were reduced for tall and increased for short persons. Slight adjustments were made to give a smooth gradation of values. The ranges were then made to correspond to ± 1 S.D. of the weights at a given height. These ranges were then arbitrarily divided, at a much later date, into three overlapping ranges to be used for "small", "medium" and "large" frame persons.

No method was established, however, for classifying persons according to their body frames. A Joint FAO/WHO Expert Committee on Nutrition (1955) has commented that such tables are "..useless, or worse, in the absence of an agreed and reliable system of classifying body types."

The figures finally derived for the Metropolitan Life tables may have been quite far removed from those found on the original individuals. Moreover, the individuals measured were of higher educational and economic status than average, and were probably not representative of the general population of the United States.

The Pryor Width-Weight tables (Pryor and Stolz, 1933) have occasionally been used in screening for obesity. These tables use the bi-iliac and lateral thoracic diameters to avoid the subjective classification into body build. The value of these tables is uncertain, however, since they are based on samples of small size.

In 1952, the Equitable Life Assurance Society produced a new set of height-weight standards for the United States. The sample studied consisted of 53,008 men and 13,943 women between

the ages of 10 and 64 years, who had purchased ordinary insurance policies from this society in 1940. When compared with the 1912 standards, the new average weights are higher for men and lower for women. Details of this study are not available to evaluate the usefulness of the standards.

In Great Britain, Kemsley (1950;1952) used data collected by the Ministry of Food in 1943. He produced tables of mean weights for 19 age groups. The weight values were also given at the 1st, 5th, 95th and 99th centiles for five heights. The values were based, however, on a small sample of three to four thousand, which decreased their reliability.

The most recent, and what may become the most useful of the height-weight tables comes from Canada (Pett, 1955 a,b,c). The figures are based on actual measurement of approximately 22,000 persons, representative of the Canadian population in 1953. Only average weights for each sex, height and age group have so far been published, but percentile ranges may become available in the near future. The Canadian figures provide a standard "that is much clearer in its derivation and significance for the population than any previous figures"(Pett, 1955c). Like the 1952 tables of the Equitable Life Assurance Society, the Canadian average weights are higher for men and lower for women than are those of the Metropolitan Life tables.

Value of Early Detection

Previous chapters have discussed the part played by obesity in the development of diabetes mellitus and essential hypertension. Maintaining normal weight is probably the most useful

preventive measure in delaying the onset of these conditions.

There is less agreement, however, on the relationship between obesity and heart disease. Keys (1955) comments that "overweight, per se, except when it is of extreme degree, is not a primary cause of coronary disease and the myocardial disorder to which it gives rise..Dieting to lose weight is highly desirable in true obesity, but a mere reduction in general food consumption is not likely to accomplish reduction of our heart disease mortality to the level prevailing in those countries with the most favorable records in this regard."

James and Hilleboe (1955) found that the frequency of overweight (defined as more than 20 per cent above the Metropolitan Life Insurance standards) among cases of coronary heart disease was similar to its prevalence among those who had no heart disease. On the other hand, Master and Jaffe (1955) found that obesity was more common in persons with coronary heart disease than in persons without coronary disease. The authors do not conclude, however, that there is necessarily an etiological relationship between the two conditions.

There is fairly considerable, though not completely convincing evidence that obesity plays a part in hastening the onset of such conditions as osteoarthritis and gall bladder disease (Pemberton, 1955; Bowser et al.,1953). Even ignoring this evidence, however, the life insurance experience, showing a considerable increase in mortality with overweight (Breslow, 1951) may suffice to justify referring obese persons for medical care.

Problems of Weight Reduction

Present day methods of weight reduction are ineffective with many individuals (Mayer, 1955). A recent Annotation (1953) suggests that "the neglect of the psychological component appears to be the principal reason for the frequent failure to reduce, and, equally important, to maintain the reduced weight level."

Bruch (1952) suggests that a rigidly enforced weight reduction program may precipitate serious mental illness. Moreover, a persisting obesity in some individuals may serve as a wall against a serious mental disturbance and may be preferable as a lesser evil (Annotation, 1953).

Mayer (1955) has recently suggested, however, that an undue emphasis may have been placed on studying the psychological aspects of obesity, with inadequate study of the physiological aspects. In support of this suggestion, Bowser et al. (1953) describe the failure of a weight reducing program, even when attention was paid to the psychological aspects.

In brief, there is much evidence that obesity is a harmful condition which is accompanied by a considerable increase in mortality. However, there is no convincing evidence that weight reducing programs are effective in reducing the prevalence of obesity or benefit the health conditions of the population.

Procedure in the Baltimore Clinic

The Metropolitan Life Insurance Company (1942;1943) tables were used, without classifying persons, however, into body build since this procedure was arbitrary and subjective. The

central figure was taken at each height for persons of medium frame. The screenee was classified as "moderately obese" when his weight was 20-29 per cent above this figure, and as "severely obese" when 30 per cent or more above this figure.

Advised to visit their family physician were:

1. All moderately obese cases with a family history of diabetes mellitus. (Moderately obese persons without the family history were advised of their overweight and allowed to decide for themselves whether or not to seek treatment).
2. All severely obese cases.

The referrals were not followed up to determine the proportion who visited their physicians. It was believed that this proportion would be small.

Results of the Baltimore Clinic

Table 1 - Results of Screening for Obesity, by Age and Sex of Persons Screened.

Screening result	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YEARS				
		2/ Total	<30	30-44	45-64	65+	2/ Total	<30	30-44	45-64	65+
Number screened	2021	922	173	402	288	57	1097	265	432	315	74
Normal	1502	668	148	294	186	39	833	235	339	199	53
Moderate obesity	227	132	16	58	51	6	94	11	31	39	9
Severe obesity	292	122	9	50	51	12	170	19	62	77	12
Rate per 1000 persons screened.											
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Normal	742	725	855	731	646	684	759	887	785	632	716
Moderate obesity	112	143	92	144	177	105	86	42	72	124	122
Severe obesity	145	132	52	124	177	211	155	72	144	244	162

1/ Includes unknown age and sex. 2/ Includes unknown age.

Table 1 of this chapter shows the results of screening for obesity.

Age. In general the prevalence of obesity increased with age until 65 years. A downward trend in prevalence occurred in those 65 years and over, although the small numbers involved prevented this trend from being significant.

It is interesting, although not shown in the tables, that 59 per cent of the moderately obese and 36 per cent of the severely obese screenees did not mark on the questionary that they were aware of their overweight.

Sex. The proportion of severely obese persons tended to be greater in women in each age group. However, the prevalence of obesity of any degree was not significantly different in the two sexes.

Table 2 - Results of Screening for Obesity, by Race and Sex of Persons Screened.

Screening Results	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	2021	1569	734	835	433	177	256
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Normal	743	758	722	790	698	751	660
Moderate obesity	112	107	146	73	118	107	125
Severe obesity	145	135	132	137	185	141	215

1/ Includes unknown race and sex.

Race. These results are shown in Table 2. The male screenees of each race showed no significant difference in their prevalence of obesity. Non-white women had a significantly greater prevalence of obesity than white women.

Table 3 - Comparison of Obesity Results and One Hour Blood Sugar, by Age and Sex of Persons Screened.

Degree of obesity, by one hour blood sugar result	BOTH SEXES, BY AGE IN YEARS					ALL AGES	
	1/ Total	<30	30-44	45-64	65+	1/ Males	1/ Females
Total screened	1993	435	825	589	128	916	1077
Above 160 mg.%	61	6	15	31	9	27	34
% above 160 mg.%	3.1	1.4	1.8	5.3	7.0	2.9	3.2
Non obese	1484	380	627	379	90	665	819
Above 160 mg.%	40	6	11	16	7	16	24
% above 160 mg.%	2.7	1.6	1.8	4.2	7.8	2.4	2.9
Moderately obese	223	27	89	87	15	131	92
Above 160 mg.%	6	0	2	3	1	4	2
% above 160 mg.%	2.7	0	2.2	3.4	6.7	3.1	2.2
Severely obese	286	28	112	123	23	120	166
Above 160 mg.%	15	0	2	12	1	7	8
% above 160 mg.%	5.2	0	1.8	9.8	4.3	5.8	4.8

1/ Includes unknown age.
Note: Only those who took both tests are tabulated.

Table 4 - Comparison of Obesity Results and Urine Sugar Test, by Age and Sex of Persons Screened.

Degree of obesity, by one hour urine sugar result	BOTH SEXES, BY AGE IN YEARS					ALL AGES	
	1/ Total	<30	30-44	45-64	65+	1/ Males	1/ Females
Total screened	1937	405	795	594	130	919	1018
Urine positive	57	13	14	23	7	34	23
% urine positive	2.9	3.2	1.8	3.9	5.4	3.7	2.3
Non obese	1435	354	602	379	92	667	768
Urine positive	44	12	13	12	7	26	18
% urine positive	3.1	3.4	2.2	3.2	7.6	3.9	2.3
Moderately obese	220	26	86	89	14	130	90
Urine positive	4	1	0	3	0	4	0
% urine positive	1.8	3.8	0	3.4	0	3.1	0
Severely obese	282	25	107	126	24	122	160
Urine positive	9	0	1	8	0	4	5
% urine positive	3.2	0	0.9	6.3	0	3.3	3.1

1/ Includes unknown age.
Note: Only those who took both tests are tabulated.

Obesity and Tests for Diabetes

Tables 3 and 4 show a tendency for the severely obese screenees between 45 and 64 years to have more positive results in the urine and blood sugar tests. However, a large proportion of the positive tests occurred in the non obese screenees. It would not seem wise, therefore, to carry out these tests only on overweight subjects.

Obesity and Blood Pressure Readings

Table 5 - Comparison of Obesity and Blood Pressure Results, by Age and Sex of Persons Screened.

Degree of obesity, by blood pressure result	BOTH SEXES, BY AGE IN YRS					ALL AGES	
	1/ Total	<30	30-44	45-64	65+	2/ Males	2/ Females
Total screened	2018	438	834	600	131	921	1095
Abnormal B.P.	149	3	37	85	22	65	84
% abnormal B.P.	7.4	0.7	4.4	14.2	16.8	7.1	7.7
Non obese	1499	383	633	382	92	667	831
Abnormal B.P.	80	3	17	43	16	34	46
% abnormal B.P.	5.3	0.8	2.7	11.3	17.4	5.1	5.5
Moderately obese	227	27	89	90	15	132	94
Abnormal B.P.	20	0	3	12	4	12	8
% abnormal B.P.	8.8	0	3.4	13.3	26.7	9.1	8.5
Severely obese	292	28	112	128	24	122	170
Abnormal B.P.	49	0	17	30	2	19	30
% abnormal B.P.	16.8	0	15.2	23.4	8.3	15.6	17.6

1/ Includes unknown age and sex. 2/ Includes unknown age.
Note: Only those who took both tests are tabulated.

Table 5 shows that the severely obese screenees between 30 and 64 years had a significantly higher proportion of abnormal blood pressure readings. This increase with obesity did not appear to occur below 30 and above 64 years.

These results confirm the frequently stated finding that hypertension is more prevalent in obese persons. The absence of this association in the screenees over 64 years may have

been due to the possibility that obese hypertensives rarely reach the age of 65 years, or that obese hypertensives over 64 did not attend the screening clinic for various reasons.

Obesity and Electrocardiograph Readings

With increasing obesity, a larger proportion of male screenees had abnormal electrocardiographic tracings. No definite increase occurred in the women, however. The severely obese men of 45-64 years had a significantly greater proportion of abnormal tracings than had the corresponding women. These changes are shown in Table 6.

Table 6 - Comparison of Obesity and ECG Results, by Age and Sex of Persons Screened.

Degree of obesity, by ECG result	TOTAL 1/	MALES, BY AGE IN YEARS 2/					FEMALES, BY AGE IN YEARS 2/				
		Total	<30	30-44	45-64	65+	Total	<30	30-44	45-64	65+
Total screened	2018	921	173	402	287	57	1095	265	432	313	74
Abnor. ECG	214	101	1	26	52	22	113	11	21	51	27
% abnor.ECG	10.6	11.0	0.6	6.5	18.1	38.6	10.3	4.2	4.9	16.3	36.5
Non obese	1499	667	148	294	185	39	831	235	339	197	53
Abnor. ECG	131	55	1	15	27	12	76	9	16	32	17
% abnor.ECG	8.7	8.2	0.7	5.1	14.6	30.8	9.1	3.8	4.7	16.2	32.1
Mod. obese	231	133	16	58	52	6	97	11	31	42	9
Abnor. ECG	31	15	0	5	7	3	16	1	2	8	4
% abnor.ECG	13.4	11.3	0	8.6	13.5	50.0	16.5	9.1	6.5	19.0	44.4
Severely ob.	288	121	9	50	50	12	167	19	62	74	12
Abnor. ECG	52	31	0	6	18	7	21	1	3	11	6
% abnor.ECG	18.1	25.6	0	12.0	36.0	58.3	12.6	5.3	4.8	14.9	50.0

1/ Includes unknown age and sex. 2/ Includes unknown age.
Note: Only those who took both tests are tabulated. ECG's read as "doubtful" not included in the abnormal ECG's.

There may be some relationship, either direct or indirect, between the prevalence of abnormal ECGs and the degree of obesity in men. This relationship is not shown in the women screenees, although their numbers are not sufficiently large to exclude this finding.

Obesity and Answers to Questionary

The relationship was studied between the presence of obesity and the answers to Questions 3 and 4 of the Self-Screening Questionary. These questions asked about discomfort on exertion, and about the need to stop to rest while walking. Although the proportion of positive answers increased with increasing obesity, the relationship was not a close or significant one.

Obesity and Presence of Albuminuria

Since obesity might affect the cardiovascular-renal system, the relationship was studied between obesity and the test results for albuminuria. There was little or no increase in the proportion of positive tests with increasing obesity.

Summary

Much work must still be done to determine the value of a screening test for obesity. While many screenees will be found to be overweight, even when the screening level is quite far from the "normal", it seems likely that many overweight screenees will not benefit from screening until weight reduction methods are more effective.

REFERENCES

Annotation.(1953). Nutr.Rev. 2:255.

Berry,W.T.C.,and Nash,F.A. (1955). Tubercle 36:164.

Bowser,L.J.,et al. (1953). J.Amer.Diet.Ass. 29:1193.*

Breslow,L. (1951) In Rosenau's Preventive Medicine and Hygiene. K.F.Maxcy,Editor. Appleton-Century-Crofts,Inc. 7th Ed.

Brozek,J.,and Keys,A. (1950). Nutr.Abstr.Rev. 20:247.

Bruch,H. (1952). Psychosomatic Med. 14:337.*

Edwards,D.A.W.,et al. (1955). Brit.J.Nutr. 9:133.

Hammond,W.H. (1955). Brit.J.Prev.Soc.Med. 9:201.

James,G.,and Hilleboe,H.E. (1955). Am.J.Public Health 45:140.

Joint FAO/WHO Expert Committee on Nutrition (1955). W.H.O. Technical Report Series No.97. Fourth Report. WHO, Geneva.

Kemsley,W.F.F. (1952). Am.Eugenics 16:316.*

Kemsley,W.F.F. (1950). Am.Eugenics 15:161.*

Keys,A. (1955). J.Chronic Dis. 1:456.

Keys,A.,and Brozek,J. (1953). Phys.Rev. 33:245.

Master,A.M.,and Jaffe,H.L. (1955). J.Amer.Ger.Soc. 3:299.

Mayer,J. (1955). Nutr.Abstr.Rev. 25:597.

Metropolitan Life Insurance Co. (1942). Statistical Bull. 23:6.

Metropolitan Life Insurance Co. (1943). Statistical Bull. 24:6.

Murphy,R. (1955). Diabetes 4:154.

Pemberton,H.S. (1955). The Practitioner 175:5.

Pett,L.B. (1955a). Canad.M.A.J. 72:12.

Pett,L.B. (1955b). Editorial. Canad.J.Public Health 46:250.

Pett,L.B. (1955c). Am.J.Public Health 45:862.

Pryor,H.B.,and Stolz,H.R. (1933). J.Pediat. 3:608.

*Original articles not read.

SCREENING FOR PROTEINURIA

Section C
Chapter 6

SCREENING FOR PROTEINURIA

Proteinuria may occur in a number of pathological conditions, which involve the urinary tract or other parts of the body. It can also be present, however, when no pathological conditions can be discovered.

Postural Proteinuria

Postural proteinuria or orthostatic albuminuria are terms usually used when temporary proteinuria occurs with the subject in the upright position, and disappears on reclining. Miller (1952) quotes the condition as occurring in 5-16 per cent of young individuals. Fishberg (1954) gives the prevalence as around 5 per cent in young adults, with the majority of cases between six and thirty years.

According to Fishberg (1954) the urine is protein free before the subject gets out of bed. The proteinuria quickly reaches a maximum during the early morning hours, and diminishes during the day. Hyaline and granular cases, and occasional red blood cells are often found. Fishberg (1954), Lyall (1953), and Conybeare and Mann (1952) all agree that postural proteinuria has no influence on the subsequent state of the kidneys, and that no treatment is required.

However, intermittent proteinuria is not completely harmless. Daley (1942) has reviewed the Medico-Actuarial Committee Report of 1938 on the mortality statistics of persons showing proteinuria. Persons with protein in two-thirds or less of the specimens examined (which presumably included most of the cases of postural proteinuria) had a mortality that was 123 per cent

of the expected. Those showing proteinuria more often than two-thirds of the time had mortalities ranging from 165 to 289 per cent of the expected. The higher figure applied to persons with the larger excretions of protein.

A recent Annotation (1955) confirms that the "immediate prognosis in cases of orthostatic proteinuria is good, and it is certain that this anomaly does not indicate rapidly progressive renal disease. Insurance statistics suggest, however, that the mortality over many years is significantly greater than in control 'good-life' cases."

In summary, the early detection of cases of postural proteinuria will not benefit these individuals. There is no treatment for the condition, and there is no agreement that the condition is harmful. In screening, the condition can be excluded to a considerable extent by asking each positive screenee to return with a specimen of urine passed immediately on rising in the morning. No referral is necessary when this second specimen is negative for protein.

Albuminuria in Pathological Conditions

According to Miller (1952) heart failure is the most common pathological condition resulting in proteinuria. This statement may be more true, however, of patients seen by clinicians than of the healthier group of persons attending a screening clinic.

Mason and Harrison (1954) emphasise that orthostatic and febrile proteinuria must be eliminated before the conclusion is drawn that renal disease may be present. The presence of leucocytes, erythrocytes and cellular and granular casts, Mason and

Harrison suggest, is indicative of active renal disease. Conybeare and Mann (1952) also affirm that the presence of more than a few red blood cells or casts on several occasions is strong evidence of an active renal lesion.

The amount of protein in the urine is not well correlated with the activity of the disease process (Mason and Harrison, 1954). The amount of protein lost in the urine in postural proteinuria may occasionally be as great as in severe glomerulonephritis (Annotation, 1955). Although most proteinurias of slight degree are benign, there are a small number which represent renal damage and may have grave prognostic significance (Bodansky, 1952). However, Diehl and McKinlay (1935) have shown that a one plus albuminuria is much more likely to be transient than a four plus result.

Benefits of Early Detection

"It seems highly probable," states Fishberg (1954), "that proteinuria, itself due to renal damage, may in turn injure the kidney." He gives the following reasons for this belief:

1. There is some evidence that readsorption and storage of protein may damage the tubular cells.
2. The coagulation of proteins which have passed through damaged glomerular capillaries may be concerned in the production of glomerular hyalinisation.
3. The tubules may become blocked by protein casts which are precipitated as the water is reabsorbed in the distal segments of the nephrons.
4. The increased viscosity due to high protein content may hamper the urinary flow down the tubule.

Fishberg apparently does not believe that postural proteinuria can produce this damage.

In brief there is little convincing evidence that the discovery and treatment of cases of proteinuria will help the individual or the community.

Results of Other Programs

Diehl and McKinlay (1935) found that 5.3 per cent of 20,000 students at the University of Minnesota had positive tests for albuminuria. 0.6 per cent (11.8 per cent of the positives) were found to have persistent proteinuria. 0.3 per cent of the total students examined had evidence which pointed to renal disease.

D'Alonzo and Rodgers (1955) described the findings of the routine pre-employment and annual physical examinations of 27,718 men between 18 and 65 years of age. Of 26,205 urine specimens, 1,030 were positive for protein. Kidney conditions were diagnosed as being present in 54 persons, of whom 18 had nephritis. 239 persons had other conditions of the genito-urinary tract which might produce proteinuria.

The results of the San Jose multiple screening clinic (American Medical Association, 1955) were almost as unsatisfactory. Of 945 persons screened, 21 gave positive tests for proteinuria. Renal disease was confirmed in 4, of whom 2 were previously unknown cases.

More satisfactory were the Indianapolis screening results (Carroll, Kurlander and Nester, 1954) obtained from a negro population. The 5,701 screenees produced 125 positives. 57 per cent of the positive screenees were confirmed positive when retested. The family physicians confirmed that 33 persons had

proteinuria, ten of these being previously unknown.

The screening of 3,988 dock workers in San Francisco (Weinerman et al.,1952) produced 92 positive tests. 35 were confirmed to have proteinuria on diagnostic examination, 16 being previously unknown.

Procedure in the Baltimore Clinic

The sulfosalicylic acid test is not falsely positive in the presence of urates and resins. The test has the advantage that no heat or caustic solutions are used. Kolmer, Spaulding and Robinson (1951) state that it is "quite sensitive and highly recommended for routine work."

The test can be conveniently carried out using proprietary tablets of sulfosalicylic acid*, and appears to be the most useful in screening for albuminuria. This was the test used in the Baltimore clinic.

The test was not carried out on menstruating women, on whom it is usually positive. Pregnant women were included in the test, however, since some women with positive tests would have developed the proteinuria since their last medical examination.

A result of one plus or more was regarded as positive. Trace results were ignored. Positive screenees were invited to return for a repeat test. Few accepted this invitation, however. When the retest was positive, or when the screening test was positive and no retest was carried out, the screenee was referred to the family physician.

The information obtained from the physicians some weeks later was of limited value. Often the physician had not completed the diagnostic examination, and was able to report only

*Bumintest tablets, Ames Co. (London) Inc.

that proteinuria was present or absent when the screenee visited his surgery. Information is therefore not complete on the number of new cases of renal disease discovered by the procedure.

Results of the Baltimore Clinic

Table 1 - Results of Screening for Proteinuria and of Follow-up with the Physician, by Age and by Sex of Persons Screened.

Screening and follow-up results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YRS				
		2/ Total	<30	30-44	45-64	65+	2/ Total	<30	30-44	45-64	65+
Number screened	1946	922	173	404	287	56	1021	232	394	310	74
Negative	1858	893	169	398	274	50	962	213	379	294	66
Positive	88	29	4	6	13	6	59	19	15	16	8
Kidney disease previously unknown	79	27	4	6	12	5	52	18	13	12	8
Follow-up with MD. completed	37	15	0	3	7	5	22	5	2	9	6
Proteinuria confirmed	8	3	0	1	0	2	5	1	0	2	2
Previously unknown to MD	4	1	0	0	0	1	3	1	0	2	0
Rate per 1000 persons screened											
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Negative	955	969	977	985	955	893	942	918	962	948	892
Positive	45	31	23	15	45	107	58	82	38	52	108
Kidney disease previously unknown	41	29	23	15	42	89	51	78	33	39	108
Follow-up with MD.. completed	19	16	0	7	24	89	22	22	5	29	81
Proteinuria confirmed	4	3	0	2	0	18	5	4	0	6	27
Previously unknown to MD	2	1	0	0	0	18	3	4	0	6	0

1/ Includes unknown age and sex. 2/ Includes unknown age.

The results of the screening procedure and follow-up pro-

cedure for proteinuria are shown in Table 1 of this chapter. In the discussion which follows, it will be assumed that a positive screening test indicated that proteinuria was present when the screenee visited the clinic. The simplicity and accuracy of the test probably justified this assumption.

Age. Both sexes under 30 years had a greater prevalence of proteinuria than those between 30 and 44 years. This greater prevalence was much more marked in women. It seems reasonable to believe that much of the increase in women under 30 years was due to postural proteinuria. The prevalence of proteinuria in both sexes was found to be 4.5 per cent, closely corresponding to the figure of 5 per cent which Fishberg (1954) gives as the prevalence of postural proteinuria in young adults.

In both sexes over the age of 44 years, the prevalence of proteinuria increased with age.

Sex. The prevalence of proteinuria was significantly greater in women under 45 than in men in the same age group. There was no significant difference in the two sexes when over 44 years.

Success of Follow-up. The older the screenee, the more likely he was to follow-up on the positive screening test. This finding was apparent in both sexes.

Confirmation of Screening Result. 22 per cent of those screenees who visited their physician were confirmed to have proteinuria. In the group of persons who visited their physician, the ratio of about four false positives to every true positive must be regarded as unsatisfactory.

Previously unknown cases. All of the 79 referrals were previously

unaware of any kidney disease, according to their answers to the screening questionary. However, only four of the eight confirmed cases of proteinuria were previously unknown to the physicians' follow-up forms.

Racial differences. Table 2 shows that there was a significantly greater proportion of non-white screenees with positive screening results than of white screenees. The racial differences occurred in both sexes.

Tavle 2 - Results of Screening for Proteinuria by Race and Sex of Persons Screened.

Screening result	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	1946	1509	734	775	417	177	240
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Negative	955	964	974	955	918	944	900
Positive	45	36	26	45	82	56	100

1/ Includes unknown race and sex.

Summary and Conclusions

1. When screening for proteinuria, a considerable proportion of positive screenees in the young adult group, especially women, will have postural proteinuria. No treatment is needed for this condition.
2. The high prevalence groups are persons under 35 and those 65 years and over, women, and non-white persons.
3. The high false:true positive ratio, and the low yield of previously unknown cases obtained in the Baltimore clinic

suggest that screening for proteinuria is of little benefit to the community.

4. When screening for proteinuria is carried out as a research project, the testing of a specimen taken immediately after the positive screenee gets out of bed may help to remove most cases of postural proteinuria.

Boyden, E.A., Spaulding, C.L., and Robinson, C.W. (1952). Appa
Laboratory Technic. Appleton-Century-Crofts, Inc. 2nd Ed.

[illegible] (1953). In Medicine. Saltzer, Gerhart, J.G., and
Williams, W. Macmillan and Gould. Vol.2.

[illegible] (1954). Integration of Intern
Medicine. The Winchester Div., Olin-Mathieson [illegible]

[illegible] (1952). Textbook of Clinical P
Williams and Wilkins Co., Baltimore. P 955.

[illegible] (1953). [illegible]

Personal Communications.

[illegible] (1952). [illegible]

REFERENCES

American Medical Association (1955). A Study of Multiple Screening. Council on Medical Service.

Annotation (1955). The Lancet 1:958.

Bodansky,O. (1952). Biochemistry of Disease. The MacMillan Co., New York. 2nd Ed. P 208.

Carroll,B.E., Kurlander,A.B., and Nester, H.G. (1954). Public Health Rep. 69:1180.

Conybeare,J., and Mann.W.N. (1952). Textbook of Medicine. The Williams and Wilkins Co., Baltimore.

Daley,R.M. (1942). Urol.Cutan.Rev. 46:13.*

D'Alonzo,C.A., and Rogers,S.M.(1955). Indust.Med.Surg.24:75.

Diehl,H.S., and McKinlay,C.A. (1935). The Kidney in Health and Disease. Editors, Berglund,H. and Medes,G. Lea and Febiger, Philadelphia. Chapter 28.

Fishberg,A.M. (1954). Hypertension and Nephritis. Lea and Febiger, Philadelphia.

Kolmer,J.A.,Spaulding,E.H.,and Robinson,H.W. (1951). Approved Laboratory Technic. Appleton-Century-Crofts, Inc. 2nd Ed.

Lyall,A. (1953), in Medicine. Editors, Garland,H.G., and Phillips,W. MacMillan and Co.Ltd. Vol.2.

Mason,M.F.,and Harrison,T.R. (1954). Principles of Internal Medicine. The Blackiston Div.,McGraw-Hill Book Co. 2nd Ed.

Miller,S.E. (1952). Textbook of Clinical Pathology. The Williams and Wilkins Co., Baltimore. P 959.

Weinerman,E.R.,Breslow,L.,Belloc,N.B.,Waybur,A. and Milmore,B.K. (1952). Amer.J.Public Health 42:1552.

General References.

Lippman,R.W. (1952). Urine and the Urinary Sediment. Charles C. Thomas, Publisher.

Todd,J.C.,Sanford,A.H.,and Wells,B.B. (1953). Clinical Diagnosis by Laboratory Methods. The W.B.Saunders Co. 12th Ed.

*Original article not read.

SCREENING FOR TUBERCULOSIS

Section C
Chapter 7

SCREENING FOR TUBERCULOSIS

The literature on this subject is now vast. Intensive surveys in Great Britain and the United States have involved much larger numbers of persons than the Baltimore clinic. It would therefore be out of place here to discuss this subject in detail, or to base the discussion on the small number of screenees attending the clinic.

The fact is now well established that most cases of early pulmonary tuberculosis show radiological changes before the physical examination is abnormal. It is also well accepted that the miniature film is satisfactory in detecting these radiological changes (Clayson et al.,1955).

Since tuberculosis is an infectious disease there are several principles involved in screening for tuberculosis which are not so important in screening for non-infectious conditions. While a program for diabetes detection might be regarded as successful if half the total unknown cases were discovered, a tuberculosis screening program must detect most unknown cases to be successful. As long as hidden cases of tuberculosis continue to spread the disease, the community has not benefitted completely from the screening program.

The time element is important in mass miniature radiography. A slow-tempo program which discovers only a small fraction of infectious cases in a given year makes little or no dent in the amount of infection in the community. Finding ten per cent of the cases per year over a nine year period, for example, is not as effective as finding ninety per cent of

the cases in one year (Perkins, 1952).

Apart from rare dissenting articles, the fact has been well accepted in the United States that intensive, high tempo programs involving the great bulk of the population are necessary for the rapid control of tuberculosis. This opinion is less widely held in Great Britain. The Rhondda Fach Survey showed that the amount of home visiting must be doubled to raise the percentage x-rayed from 75 to 90 per cent (Cochrane, 1954). Stein (1954) and others suggest that the money involved in expensive surveys might be used to more effect by improving housing conditions.

Benefits of Early Discovery

Chest x-ray surveys detect cases of tuberculosis, who are placed under treatment and cease transmitting their disease to others, sooner than ifthe surveys had not been conducted. It does seem likely, therefore, that the community benefits from such surveys. There has been frequent argument, however, on when active cases found by surveys would be diagnosed in the absence of a mass radiography program, and on whether such cases benefit by the early discovery. Clearly a study would not be morally acceptable which used as controls persons who attended the survey, were positive for tuberculosis, but who were allowed to go untreated until discovered by some other method.

Anderson and others (1954) attempt to overcome this difficulty. They compare a group of 135 active cases of pulmonary tuberculosis discovered in a Minneapolis survey in 1947 with a group of active pulmonary cases reported to the city health department just prior to the survey.

The four year follow-up for the survey roughly covers the years 1948 to 1952, while for the non survey cases is 1947 to 1951. The difference in mortality between the two groups is believed too great to be accounted for by the slightly improved therapy given to the survey group.

Thirty-seven per cent of the survey cases were discovered in the minimal stage, compared with 33 per cent of the non survey cases. Sixteen per cent of the survey cases were far advanced, compared with 33 per cent of the non survey cases. Using the life table method to study the survival of these two groups of cases over the four year period, the authors estimate that the chance of a survey case dying within four years is one out of ten, whereas the chance of a non survey case dying is one out of three.

Anderson et al. (1954) conclude that the differences observed in the outcome of tuberculosis in the two groups of cases are to some extent a reflection of the value of early case discovery through intensive tuberculosis case-finding efforts.

This study is not likely to end the argument on this subject. There is no mention of the social status, income group or attitude to the infection of the individuals in the two groups. It is possible, therefore, that the 135 survey cases, even if discovered at a later stage, might still have an improved survival rate over the control group in the study. The very separation into persons who attended an x-ray survey and

those who would not necessarily attend might produce groups with differing survival rates.

Hospital Admission Chest X-Ray Programs

Routine hospital admission chest x-ray programs produce a greater yield of active cases per 1,000 persons x-rayed than mass surveys of the general population.

Siegal, Pluckett and Locke (1955) compare the results of the hospital admission survey and the mass survey programs in New York State during 1947 to 1953. Of previously unreported cases initially classified as definite or suspected tuberculosis, 69 per cent were followed-up to produce 1.7 new active cases per 1,000 persons x-rayed in the hospital program. If those not followed-up had the same proportion of cases, 100 per cent follow-up would have produced 2.4 new active cases per 1,000 x-rayed. The corresponding figures for the mass radiography program were 1.3 newly discovered active cases per 1,000 x-rayed with 71 per cent follow-up, and 1.8 per 1,000 with 100 per cent follow-up.

The increased yield of newly discovered active cases in the hospital program does not tell the full story, however. The following table shows the percentage distribution, according to apparent stage of disease, of persons classified on screening as probably active tuberculosis.

PROGRAM	MINIMAL	MODERATELY ADVANCED	FAR ADVANCED
Hospital Admission	37.0	40.8	22.1
Mass Radiography	51.5	38.5	9.9

Proportionately, twice as many hospital admission cases are far advanced than are the mass survey cases. Moreover, a larger proportion of the hospital admission cases has presumably spread the infection to other members of the community than the proportion of mass radiography cases.

Jacobson and Adler (1954) quote the costs of similar programs in Los Angeles County during 1951-52. Approximately $180 is spent to find a previously unknown active case in the hospital program. The equivalent cost in the mass survey program is $808. They suggest:"..it would appear that a widespread hospital admission chest program should be the greatest and most economical means for finding tuberculosis, considering that approximately 16,000,000 people are admitted each year to general hospitals throughout the United States. These patients constitute a large segment of the population readily available for examination, rapid diagnosis and follow-up.

"A large proportion of those admitted to general hospitals are in the older age groups, in which it has been shown that the highest prevalence of tuberculosis exists."

It is thus apparent that one program will frequently reach persons not usually involved in the other. The cost figures given are not necessarily an argument in favour of concentrating on one type of program alone, since discovering a far advanced case is not as beneficial to the community as detecting an early case.

In summary, the hospital admission and mass survey programs

are complementary procedures, and not programs by which concentration on one would justify the exclusion of the other.

Results with Different Economic Groups

Anderson, Enterline and Turner (1954) compare their findings in the highest and lowest economic groups in a county in Ohio. The tuberculosis death rate was twenty times greater in the lowest economic group, the prevalence of cases known to the health department was three times greater, and the estimated prevalence of previously unknown active cases was one and one-half times greater than in the highest economic level.

These investigators suggest that tuberculosis may be more chronic in the higher economic groups, less frequently terminating in death and less frequently appearing to be of sufficient clinical significance to warrant a notification to the health department. They therefore conclude that the higher economic groups should not be overlooked in mass miniature radiography programs.

Survey Needs of Different Racial Groups

An intensive community-wide tuberculosis survey in Georgia in 1946 (Comstock and Burke, 1951) was followed by continuous case-finding and reporting activities, and by a second survey in 1950 (Comstock and Sartwell, 1955). These studies show that the incidence of new cases is much greater in negroes, and suggest that more frequent and more intensive surveys are necessary for this racial group.

Error in Miniature Film Interpretation

Birkelo et al. (1947) showed a variation in interpreting the same films by different observers, and a slightly smaller variation when the same films were interpreted by the same reader at different times. Their study has been confirmed by investigators in Great Britain and elsewhere (Groth-Petersen et al.,1952). There is now fairly general acceptance that a single observer, attempting to pick out cases of pulmonary tuberculosis from a largely normal population, is likely to overlook about 25 to 30 per cent of the abnormal films (Zwerling et al.,1951; Editorial, 1955).

"Boredom and the hypnotic effect of the repeated normal" have been suggested as the major cause of this error. "Opinion can often become biased in favour of the predominant; after reading many more or less uniform films the observer becomes less inclined to remark on the slightly exceptional." (Editorial, 1955).

On the basis of these findings, Birkelo and others (1947) have recommended that all films should be read independently by at least two interpreters -- dual reading -- and that any individual read as positive by one of the readers should receive further study.

Dual Reading of Miniature Films

Representative of the studies of dual reading are those of Yerushalmy et al. (1950) in the United States and by Groth-Petersen and Moller (1955) in Denmark. These studies show that

dual reading can reduce the false negatives by one- to two-thirds, increasing the sensitivity of the method from 70 per cent (when one reading only is carried out) to 80 or 90 per cent. At the same time, however, the number of false positive results are usually doubled.

This serious loss in specificity is important, since twice the number of persons have to be recalled for further study to produce the comparatively small increase in true positives. Nevertheless, the investigators conclude that "dual reading offers an easy and inexpensive way of increasing the yield of radiographic case-finding." Indeed, a Lancet Editorial (1955) suggests: "Whatever the purpose of the examination, and whatever the size of the film, it is now apparent that the radiograph should preferably be read by two observers."

Although this recommendation has been widely followed in Denmark, it has not been followed in the United States or Great Britain. The shortage of qualified readers, and the increase in false positives have probably been the main reasons.

The reading studies have shown the unusual ability of occasional readers to give results just as satisfactory as dual reading, without the corresponding increase in false positives. There appear to be certain sources of error which can be defined so that less efficient readers can learn to avoid them (Groth-Petersen and Moller, 1955).

Use of a Second Miniature Chest Film

Stradling and Johnston (1955) suggest a method of reducing

the increase in false positives produced by dual reading. In addition to the usual postero-anterior film, they routinely took an anteroposterior lordotic view of the chest. This latter film may reveal lesions overlooked or concealed behind ribs or clavicles in the former view. These investigators found that the combination of the two views reduced by almost one-half the false positives, both in single reading and in dual reading. They believed that the initial increase in expense was more than balanced by the reduced number of full size films necessary. This method seems worthy of further investigation.

Procedure in the Baltimore Clinic

A 70 mm. postero-anterior chest film was taken routinely. A relatively large number of films were technically unsatisfactory, partly because a trained radiographer was not available during the first few weeks of the clinic. The screenee was not asked to return when this occurred, but was merely told that the film had been technically unsatisfactory.

An experienced chest physician read the miniature films. When the small film could not exclude tuberculosis, the screenee was requested to have a full-size film taken in the nearest health department chest clinic. The screenee was referred when the result was still doubtful or abnormal, or when the screenee failed to have the large film. When it was known that the screenee had visited the doctor, the physician was contacted several weeks later to determine the result of the referral.

Results of the Baltimore Clinic

The results are shown on page 189.

Table 1 - Results of 70 mm. chest x-ray in screening for tuberculosis and of follow-up with the physician, by age and sex of persons screened.

Screening and follow-up results	Total 1/ persons screened	BOTH SEXES, BY AGE IN YEARS				ALL AGES 3/	
		Under 2/ <30	30-44	45-64	65+	Male	Female
Number screened	1767	403	750	496	106	818	948
No evidence of Tbc	1733	398	739	482	102	802	930
Tbc to be ruled out	34	5	11	14	4	16	18
Tbc previously unknown	28	4	10	11	3	14	14
Follow-up with MD completed	18	4	3	6	3	4	13
Tuberculosis confirmed	2	0	1	1	0	1	1
Previously unknown to MD	2	0	1	1	0	1	1
	Rate per 1000 persons screened						
Number screened	1000	1000	1000	1000	1000	1000	1000
No evidence of Tbc	981	988	985	972	972	983	985
Tbc to be ruled out	19	12	15	28	38	17	15
Tbc previously unknown	16	10	13	22	38	17	15
Follow-up with MD completed	9	10	4	12	28	4	14
Tuberculosis confirmed	1	0	1	2	0	1	1
Previously unknown to MD	1	0	1	2	0	1	1

1/ Includes unknown age and sex. 2/ Includes unknown sex.
3/ Includes unknown age.

Age. The proportion of positive results increased with age.

Sex. There was no significant difference in the results of the two sexes.

Success of follow-up. Women with positive results were much more faithful in attending for follow-up than men.

False positive:true positive ratio. This was 13:1, based on

total referrals, and 8:1 based on those who completed the follow-up.

Previously unknown cases. Both confirmed cases were previously unknown to the physician.

Race. Table 2 shows that there is a smaller proportion of abnormal screening findings in the non-white group. The difference is not significant, however.

Table 2 - Results of 70 mm. chest x-ray in screening for tuberculosis, by race and sex of persons screened.

Screening result	Total 1/ persons screened	WHITE			NON-WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Number screened	1767	1369	652	717	383	157	226
	Rate per 1000 persons screened						
Number screened	1000	1000	1000	1000	1000	1000	1000
No evidence of Tbc.	981	980	980	979	984	981	987
Tbc to be ruled out	19	20	20	21	16	19	13

1/ Includes unknown race and sex.

Yield. The yield was approximately one previously unknown case per thousand screenees, similar to that found in much larger surveys in the United States.

Summary

While the false positive:true positive ratio is higher and the yield is lower than those of some other screening tests, the importance of the condition detected justifies the use of the 70 mm. chest x-ray in screening for tuberculosis.

Postscript on Screening for Lung Carcinoma

The possibility that mass miniature radiography might be helpful in detecting early carcinoma of the lung has been under study for some years. In many surveys the films are now read routinely for possible neoplasms in addition to tuberculosis, heart disease and other chest conditions.

The findings of the Baltimore clinic are not helpful with respect to lung carcinoma and other chest conditions, because of the relatively small number of persons involved, and will not be presented here. It might be useful, however, to summarize a few of the latest papers on this subject.

The average yield of proved cases of bronchogenic cancer is ten per 100,000 adults examined (Garland, 1955). This figure will vary with the age distribution of those attending the survey, and is greatest in males over 45 years of age (Guiss, 1955).

With present techniques, however, less than ten per cent of survey-detected lung cancer appears to be curable (Boucot and Sokoloff, 1955). This poor result is partly due to the progress of the cancer during the frequently protracted delay between the survey and the operation (Guiss, 1952).

The failure of 70 mm. films to show up bronchogenic carcinomata at an early stage is the other major reason for the lack of success (Boucot and Sokoloff, 1954). McNulty (1954) suggests that surveys must be repeated indefinitely at intervals of not more than four to six months in order to detect malignant lesions as they arise. He recommends two exposures,

one in inspiration and one in expiration, with at least two competent radiologists to interpret the films and to check themselves on cases where there is disagreement (Pendergrass, 1952).

The preliminary results of a study which incorporates most of these suggestions have recently been published (Boucot et al., 1955). The Philadelphia Pulmonary Neoplasm Research Project involves the following of 6,000 men, 45 years of age and older, by 70 mm. x-rays, both inspiratory and expiratory, taken every six months for a period of ten years.

Thirty-seven proven cases of lung cancer have been detected in the initial stage of this study. Eleven have been resected and have remained alive within the short period of follow-up. Since this study is being carried out by experienced and enthusiastic workers with a cooperative group of patients, it is likely to have better results than any mass radiography program in which carcinoma is only one of the conditions being screened. Unless these early results are not representative of those to be obtained later in the study, "the suitability of single photofluorograms for finding curable lung cancer must be questioned" (Boucot et al.,1955). Indeed, the long-hoped-for serological test for malignancy may be the only screening method likely to help in the early detection of lung carcinoma.

REFERENCES

Anderson,R.J. et al. (1954). Public Health Rep. 69:1053.

Anderson,R.J., Enterline,P.E., and Turner,C.D. (1954). Amer.Rev.Tuberc. 70:593.

Birkelo,C.C. et al. (1947). J.Amer.Med.Ass. 133:359.

Boucot,K.R. et al. (1955). J.Amer.Med.Ass. 157:440.

Boucot,K.R., and Sokoloff,M.J.(1954). Amer.Rev.Tuberc. 69:164.

Boucot,K.R., and Sokoloff,M.J. (1955). Dis.Chest 27:369.

Clayson,C. et al.(1955). Brit.J.Tub.Dis.Chest 49:81.

Cochrane,A.L. (1954). Brit.Med.Bull. 10:91.

Comstock,G.W.,and Burke,M.H. (1951). Public Health Rep.66:695.

Comstock,G.W.,and Sartwell,P.E. (1955). Amer.J.Hygiene 61:261.

Editorial (1955). The Lancet 1:1257.

Garland,L.H.(1955). Guest Editorial. J.Amer.Med.Ass.157:1499.

Gröth-Petersen,E. et al. (1952). Acta Tub.Scand. 26:13.

Groth-Petersen,E., and Moller,A.V. (1955). Bull.Wld.Hlth.Org. 12:247.

Guiss,L.W. (1955). Cancer 8:219.

Jacobson,G.,and Adler,D.C. (1954). Amer.Rev.Tuberc. 69:940.

McNulty,J.M. (1954). New England J.Med. 250:14.

Pendergrass,E.P.(1952). Proceedings of the Second National Cancer Conference. Vol.I. P 70. American Cancer Society, Inc.

Perkins,J.E. (1952). Community-Wide Chest X-ray Survey. Public Health Service Publication No.222. U.S.Govt.Printing Office, Washington,D.C. P 7.

Siegal,W., Pluckett,R.E., and Locke,B.Z. (1955). J.Amer.Med.Ass. 157:435.

Stein,L. (1954). Tubercle 25:195.

Stradling,P.,and Johnston,R.N. (1955). The Lancet 1:1247.

Yerushalmy,J. et al. (1950). Amer.Rev.Tuberc. 61:443.

Zwerling,H.B. et al. (1951). Amer.Rev.Tuberc. 64:249.

earlier art elsewhere, this subject will only be discussed briefly here.

Serological Tests

The serologic[illegible] for two types of [illegible] (Editorial, 1954):

1. The purified phospholipid antigens, such as cardiolipin
2. The crude [illegible] alcoholic-extract antigens -- the "[illegible] antigens."

SCREENING FOR SYPHILIS

Section C
Chapter 8

In general, cardiolipin [illegible] more sensitive than crude extract antigens (Editorial, 1954; Index et al., 1954) in detecting reagin earlier in the disease, and they remain positive longer in the declining phase of the serological in cases under treatment (Editorial, 1954). Although opinions differ, most serologists regard cardiolipin antigen as be more specific (Editorial, 1954; Ladon et al., 1954).

While the cardiolipin tests seem to be slightly more sensitive, it would appear that any of the well recognized serological tests carried out by an efficient laboratory will satisfactorily be screening for syphilis.

Screening in Syphilis Control

"The syphilis screening program," state Smith, Don[illegible]

SCREENING FOR SYPHILIS

Since large scale screening programs for syphilis have been carried out elsewhere, this subject will only be discussed briefly here.

Serological Tests

The serological tests for syphilis use two types of antigen (Editorial, 1954):

1. The purified phospholipid antigens, such as cardiolipin.
2. The crude ox-heart alcoholic-extract antigens -- the "lipoidal antigens."

In general, cardiolipin antigen is more sensitive than crude ox-heart antigen (Editorial, 1954; Isdoe et al.,1954). It detects reagin earlier in the disease, and the test remains positive longer in the declining phase of the serological pattern in cases under treatment (Editorial, 1954). Although opinions differ, most serologists regard cardiolipin antigen as being more specific (Editorial, 1954; Isdoe et al.,1954).

While the cardiolipin tests seem to be slightly more efficient, it would appear that any of the well recognised serological tests carried out by an efficient laboratory will be satisfactory in screening for syphilis.

Screening in Syphilis Control

"The syphilis casefinding program," state Smith, Donohye and Stuart (1955), "is an integrated machine with serologic screening as a basic component and with presurvey education and postsurvey contact investigation as necessary partners."

The Baltimore clinic did not have personnel to carry out

the recommended contact investigation of positive screenees. It was therefore realised that the serological testing alone would have little effect in the control of syphilis in this population. However, it was felt that a knowledge of the prevalence of positive serological tests would be helpful to the Commission's study.

Treponema Pallidum Immobilization Test

The TPI test reacts to a treponemal immobilizing antibody which is separate from the reagin which causes the positive serological test (Expert Committee, 1954). The TPI test can therefore be positive when the serological test is negative, and vice versa. However, the test is complicated, time consuming, expensive, and can only be carried out in a limited number of laboratories (Expert Committee, 1954). It is therefore not suitable for routine administration in a screening program (MacPherson, Ledbetter and Martens, 1955), but may be helpful if carried out on screenees with positive serological tests.

In the early stages of syphilis, the TPI test takes longer than the standard tests to become positive (Editorial, 1954). During treatment of the disease, the TPI results run parallel with those of the standard tests, but the standard test results become negative more quickly (Editorial, 1954).

Occasionally the immobilization antibody will persist in the adequately treated case of early syphilis. The immobilizing antibody will usually persist in the case of late syphilis, whether treated or untreated, although the serological test may become negative (Isdoe et al.,1954).

The TPI test is, therefore, not a guide to the need for treatment, since it is usually present in the adequately treated case of late syphilis and may frequently be absent in the case of primary or early syphilis. The latter is rarely encountered in screening in the United States. The screenee with a negative TPI test and a positive serological test is usually a "biological false positive" (BFP) with a non treponemal condition.

Leprosy, disseminated lupus erythematosus, and a large number of acute infections produce BFP reactions (Moore and Mohr, 1952). Moore and Lutz (1955) comment that "..these laboratory changes may persist for many years in the absence of any clinical disease at all. Nevertheless, clinical disease usually does develop in many chronic BFP reactors; and the phenomenon is far from innocuous.

"Indeed, the physician is no longer justified, when he has identified a chronic BFP reactor, in dismissing his patient with congratulations on the absence of syphilis. Instead, he is faced with a lengthy and detailed clinical investigation to attempt to identify the cause of the BFP reaction; and this may and usually does mean prolonged periodic observation and re-examination."

Procedure in the Baltimore Clinic

The VDRL Slide test and Eagle Slide test were carried out on all blood specimens during the first four weeks. In the last six weeks ofthe clinic the VDRL Slide test and TPI test were carried out routinely. The result of the TPI test was reported only when the serological test was positive. It was known that

a number of persons would have positive TPI tests with negative serologies. It was felt, however, that this finding would not be helpful to the family physician, since the significance of a positive TPI test alone has not yet been clarified.

Screenees with both a positive STS and TPI test were referred for clinical evaluation, unless the serological result was of low titre (1:1 or below) or the TPI result doubtful. Those individuals with positive STS and negative TPI results were not referred for further examination. Whether this decision was wise or not is difficult to say, in view of the comments of Moore and Lutz (1955). It was difficult to require those persons, at a time when much more has to be learned of such conditions, to undergo the expensive study necessary to exclude the conditions which give biological false positive results. The screening clinic should perhaps provide special diagnostic facilities for such cases.

Results of the Baltimore Clinic

The results are shown in Table 1 of this chapter.

Age. In general, the proportion of positive tests increased with age.

Sex. There were no consistent differences in the proportion of positive results between the two sexes.

Completion of follow-up procedure. Less than half of the 34 referrals were known to have gone for diagnostic examination.

False positive:true positive ratio. This was 7.5:1 when all the referrals were included, and 2.5:1 when only those who completed the follow-up were included.

Table 1 - Results of Syphilis Tests and of Follow-up with Physician, by Age and Sex of Persons Screened.

Screening and follow-up results	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YRS				
		2/ Total	<30	30-44	45-64	65+	2/ Total	<30	30-44	45-64	65+
Number screened	1949	912	171	403	280	56	1034	250	412	292	69
Negative	1896	884	170	395	262	55	1009	249	400	283	66
Doubtful or positive	* 53	28	1	8	18	1	25	1	12	9	3
Follow-up with MD. completed	14	6	0	1	4	1	8	0	3	3	2
Syphilis confirmed	4	2	0	0	2	0	2	0	1	1	0
Previously unknown to MD	1	1	0	0	1	0	0	0	0	0	0
Rate per 1000 persons screened											
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Negative	973	969	994	980	936	982	976	996	971	970	957
Doubtful or positive	27	31	6	18	46	18	24	4	29	30	43
Follow-up with MD. completed	7	7	0	2	10	18	8	0	7	10	28
Syphilis confirmed	2	2	0	0	5	0	2	0	2	3	0
Previously unknown to MD	0.5	1.1	0	0	2.6	0	0	0	0	0	0

1/ Includes unknown age and sex. 2/ Includes unknown age.
* Approximately 34 referred for diagnostic examination.

Table 2 - Results of Screening for Syphilis, by Race and Sex of Persons Screened.

Screening result	Total 1/ persons screened	WHITE			NON WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total persons screened	1949	1513	725	788	417	176	241
	Rate per 1000 persons screened						
Total persons screened	1000	1000	1000	1000	1000	1000	1000
Negative	973	988	992	985	916	875	946
Doubtful or positive	27	12	8	15	84	125	54

1/ Includes unknown race and sex.

Previously unknown cases. The screening questionary did not ask the individual if he already knew he had syphilis, since it appeared likely that few answers would be affirmative. Of the four confirmed cases of syphilis, three were previously known to their physicians.

Race. Table 2 shows that the proportion of positive results was markedly greater in both sexes of the non white race.

Discussion of Results

The yield of 0.5 previously unknown cases per 1000 screenees is unsatisfactory, and suggests that screening the general public for syphilis is not a valid procedure in the United States. However, the reliability of the follow-up findings for syphilis is probably even lower than that for other diseases. Special diagnostic facilities for the referrals would have been necessary to improve the follow-up.

Table 3 - TPI results of Screenees with Doubtful or Abnormal Serologies.

STS RESULT	Total	TPI RESULT	
		Non reactive	Reactive
Doubtful or positive	45	6	39
Doubtful	15	2	13
Positive	30	4	26

The TPI test gave some help in deciding whether or not a screenee should be regarded as a possible syphilitic. Table 3 shows that six screenees with doubtful or abnormal serological results had non reactive TPI tests, and were not referred for

diagnostic examination. Factors mentioned earlier, however, make this test unsuitable for routine administration in a screening program.

Conclusion

The inclusion of a serological test for syphilis in a multiple screening program for the general public produces too small a yield of previously unknown cases to be worth while. However, this yield may be more productive if the testing is confined to high prevalence groups -- for instance, the negro population in Baltimore -- and if the diagnostic examination is carried out by facilities with a special interest in syphilis.

REFERENCES

Editorial (1954). The Lancet 2:1165.

Expert Committee (1954). Expert Committee on Venereal Infections and Treponematoses, Subcommittee on Serologic and Laboratory Aspects. Third Report. World Health Organization Technical Report Series No. 79.

Isdoe,O. et al. (1954). Bull.Wld.Hlth.Org. 10:507.

MacPherson,D.J.,Ledbetter,R.K. and Martens,V.E. (1955). Amer.J.Clin.Path. 25:89.

Moore,J.E. and Lutz,W.B. (1955). J.Chronic Dis. 1:297.

Moore,J.E. and Mohr,C.F. (1952). Ann.Int.Med. 37:1156.

Smith,C.A.,Donohue,J.F. and Stuart,J. (1955). J.Chronic Dis. 2:480.

General Reference.

Simmons,J.S. and Gentzkow,C.J. (1955). Medical and Public Health Laboratory Methods. Lea and Febiger, Philadelphia.

VISION, HEARING AND DENTAL SCREENING TESTS

Section C
Chapter 9

SCREENING FOR IMPAIRED VISUAL ACUITY

A number of screening clinics have included tests for impaired visual acuity. The simplicity of the tests used and the high prevalence of the condition detected appear to be the main reasons for their inclusion in screening.

Sunderman and Boerner (1950) quote the normal range of visual acuity as 20/15 to 20/40. "From the age of 14 to about 25 there is probably a gradual increase of visual acuity; a very gradual decline then begins which is not attributable to disease." (Keil,1949). There is also a gradual loss of accommodative power as age advances (Luedde, 1949). Visual acuity measurements are fairly constant and change very little in the same individual from day to day without pathological cause.

In the emmetropic eye of the young individual, parallel rays of light from a distant object are focused on the retina when the eye is unaccommodated. Thus the test for distant visual acuity is mainly a test of the ability of the retinal elements to distinguish the form of the object. The rays of light from a near object, however, will not be focused on the retina unless adequate accommodation occurs. The so-called test for near visual acuity, therefore, is actually a combined test of accommodative power and of the ability of the retinal elements to distinguish the shape of the object. Both visual acuity and accommodative power diminish with age (Keil, 1949;Luedde, 1949). It would be expected, therefore, that considerable numbers of elderly persons will fail in a test for near vision

when no corrective glasses are used.

Newell and Zinn (1955) emphasize certain points which are frequently forgotten in screening for impaired visual acuity. "The eye with 20/20 vision may be described as normal, only if, in addition to the clear central vision, reading vision is adequate, with or without correcting lenses; color vision is normal; the peripheral field of vision is normal; there are no defects in that area of central vision within 30 degrees of the fixation point; and the two eyes can be used simultaneously and move normally. Lastly, if careful ophthalmoscopic examination of the disk indicates no abnormalities, the blood vessels are good, the macula is normal, and peripheral fundus shows no disease, then it is justifiable after refraction to conclude that the 20/20 vision is associated with a normal pair of eyes."

"The finding of 20/20 vision .. does not .. imply that the eyes are without disease and may not require correcting glasses" (Newell and Zinn, 1955). This fact may explain the not infrequent finding of screenees who have been required to wear glasses although the visual acuity appears to be satisfactory. Moreover, "visual acuity is no clue to the magnitude of refractive error. It is entirely possible for one individual to have 20/20 vision with a given refractive error while another can read only 20/40 vision with the same refractive error." (Newell and Zinn, 1955). A number of individuals with impaired accommodative powers will be missed, therefore, even if the screening level used is the normal value of 20/20. Raising the screening level will tend

to detect the more severe refractive errors, but will not invariably do so.

Caccamise (1954) stresses that no actual damage is ever done by not wearing corrective glasses, although symptoms of eyestrain may occur. Finding adults with impaired visual acuity does not usually result in the detection and treatment of progressive conditions. However, there is more definite evidence that the similar screening of children will detect progressive conditions which will benefit by their early discovery.

Procedure in the Baltimore Clinic

Probably the most widely used tests for visual acuity in the United States are the Snellen chart for distant vision and Jaeger type for near vision. Both of the test cards must be held at a definite distance from the eyes, and must be adequately and uniformly illuminated.

In order to economise in space, and to overcome the difficulty in standardising the illumination of the test cards, the Baltimore clinic used the American Optical Company Sight Screener. This instrument is simple to use, tests both near and distant vision, and will test each eye individually without requiring the screenee to close his eyes at any time.

The test was first carried out without corrective glasses to obtain prevalence figures. If the screenee normally wore corrective glasses, the test was repeated while he was wearing these glasses. Each eye was tested individually for near and distant vision, the test for binocular visual acuity being omitted.

The results while the screenee was wearing his usual glasses, if any, were used to determine the need for further examination. Where distant or near vision was 20/40 in each eye with corrective glasses, or below 20/40 in one eye, the screenee was told that he might benefit by a more thorough examination. (This was not advised, however, when his glasses had been prescribed within the previous year.) The screenee was informed of the actual result before he left the clinic.

No follow-up results were obtained. The experience of other screening clinics has suggested that only a small proportion of the positives seek a more thorough examination. It was also felt that the follow-up efforts should concentrate on individuals whose positive tests suggested that a progressive condition was present.

Results of the Baltimore Clinic

The results are summarized in Tables 1 and 2 of this chapter.

Age. A marked increase in the percentage of persons failing in the test occurred with each succeeding age group. In those persons 45 years of age or older, more than 90 per cent were unable to pass the test without glasses.

Sex. No significant differences occurred between the two sexes.

Race. A smaller proportion of non-white persons failed in the test without glasses.

Summary and Conclusions

When a test for visual acuity, with a screening level about 20/40 or 20/50, is included in the screening procedure, considerable numbers of screenees will fail to pass the test. The proportion of failures depends mainly on the age distribution

Table 1 - Results of Screening for Impaired Visual Acuity, by Age and Sex of Persons Screened.

Screening procedure	Total 1/ persons screened	BOTH SEXES, BY AGE IN YEARS 2/				SEX-ALL AGES 3/	
		Under <30	30-44	45-64	65+	Male	Female
Total screened without glasses	2006	437	831	598	124	919	1084
Per cent abnormal	52.8	19.0	33.0	95.0	98.4	52.2	53.2
Number screened with glasses	993	104	293	472	116	440	552
Per cent abnormal	30.8	5.8	20.1	36.9	55.2	30.7	31.0
Number screened without glasses, not subsequently screened with glasses	1013	333	538	126	8	479	532
Per cent abnormal	22.1	7.2	15.1	84.9	87.5	21.9	22.0

1/ Includes age and sex not stated. 2/ Includes sex not stated.
3/ Includes age not stated.

Table 2 - Results of Screening for Impaired Visual Acuity, by Race and Sex of Persons Screened.

Screening procedure	Total 1/ persons screened	WHITE			NON WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total screened without glasses	2006	1558	732	826	428	176	252
% abnormal	52.8	54.6	53.6	54.8	47.2	46.6	47.6
Number screened with glasses	993	824	375	449	159	60	99
% abnormal	30.8	28.8	29.6	28.1	42.1	40.0	25.0
Number screened without glasses, not subsequently screened with glasses	1013	734	357	377	269	116	153
% abnormal	22.1	19.8	19.6	19.9	27.5	27.6	50.0

1/ Includes race and sex not stated.

of the screenees.

It is likely that only a small proportion of the positive screenees will be sufficiently interested to seek further examination. Moreover, the conditions detected are usually not cured nor is their progress delayed by their early discovery. However, the test is simple, inexpensive, and may result in those persons who do follow-up on the positive result getting more pleasure out of life. The value of including such a test in the multiple screening procedure is therefore a matter of opinion.

REFERENCES

Caccamise,W.C. (1954). G.P. 10:82.

Keil,F.C. (1949) in the Eye and Its Diseases. Editor,C.Berens. The W.B.Saunders Co. 2nd Ed.

Luedde,W.H. (1949) in The Eye and Its Diseases. Editor,C.Berens. The W.B.Saunders Co. 2nd Ed.

Newell,F.W. and Zinn,S. (1955). Postgrad.Med. 17:81.

Sunderman,F.W. and Boerner,F. (1950). Normal Values in Clinical Medicine. The W.B.Saunders Co.

General References.

Best,C.H. and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. The Williams and Wilkins Co. Tenth Ed.

Adler,F.H. (1947). Gifford's Textbook of Ophthalmology. The W.B.Saunders Co. 4th Ed.

SCREENING FOR HEARING IMPAIRMENT

In recent years in the United States, many schools have routinely tested children for hearing difficulties. The proportion of correctible defects has apparently been sufficient to justify continuing these programs. Probably encouraged by this, and perhaps also by the availability of a simple, rapidly administered test, screening clinic administrators have sometimes included such a procedure in their adult screening programs.

Hearing loss in the higher frequencies is comparatively common in middle and later life (Houssay et al.,1955), and has been termed presbycusis. This gradually progressive, physiologic loss of hearing for the high-pitched tones develops as the result of arteriosclerotic and degenerative changes within the inner ear. This loss results in diminished understanding of the consonants. It is not a noticeable handicap in all elderly people, however, and even after the age of seventy years many persons are found to retain remarkably acute hearing (Best and Taylor, 1955).

The National Health Survey (U.S.Public Health Service) included a study of the hearing threshold acuity measurements of both ears of 2,002 men and 2,660 women. All had a clinical history of normal hearing for speech in both ears. The loss with age of hearing in the higher frequencies was more marked in men than in women. The mean hearing threshold acuity in the 30-39 year age group, for instance, was 21.6 decibels for men and 10.5 decibels for women. The corresponding figures for the

sixty years and over age group was 39.7 decibels for men and 22.2 decibels for women.

These results suggest that, when multiple screening includes a test for hearing impairment, one of two alternatives seems advisable:

a. The screening level should change with differences in age and sex, or
b. The screening test should not include the higher frequencies. This method was used in the Baltimore clinic.

Procedure in the Baltimore Clinic

The screening test was used, not only to determine the need for more thorough examination of the hearing, but also to obtain additional information on the prevalence of mild hearing loss which was not necessarily disabling. The test was therefore more elaborate than usual.

A sweep check frequency test was carried out on a pure tone audiometer. A "sound proof" booth was used which was not, however, able to keep out all noise in the hall. Each ear was tested at frequencies of 125,250,500,1000,2000,4000 and 8000 cycles per second, using an intensity of 20 decibels. An individual was"positive" who failed to hear two or more frequencies in one ear. The positive screenees were not infrequently found to have the hearing loss in both ears, especially when presbycusis was present.

Before they left the clinic the positive screenees had a more thorough rescreening test. This test was administered in an adjacent room which was only occasionally affected by outside noise.

The rescreening test was a hearing threshold test carried out at the same frequencies, again using a pure tone audiometer. The result was recorded as "abnormal" when the better ear had an average hearing threshold of 30 decibels or more at the frequencies of 500, 1000 and 2000 cycles per second. Persons with this result had a handicap which might be severe enough to need a hearing aid. These individuals were told of this result during their interview at the clinic. However, almost all were already aware of their hearing loss, and most appeared disinclined to seek the help of a hearing aid.

The result was recorded as "doubtful" when not in the abnormal category, but when a threshold of 25 decibels or more at two or more frequencies was found in at least one ear. Individuals with this result were classified as "potentially handicapped persons", although they would not necessarily progress into a handicapping stage of deafness. Many elderly screenees fell into this category.

Neither "abnormal" nor "doubtful" screenees were urged to visit their personal physicians. Most general practitioners in the United States do not have audiometers to confirm or deny the screening findings. To obtain adequate follow-up of the referrals, it would probably have been necessary to refer them to otologists, or to set up special diagnostic facilities.

Results of the Baltimore Clinic

The results of the tests for hearing impairment are given in Tables 3 and 4 of this chapter.

Age. The prevalence of "handicapped" persons increased markedly with age. The prevalence of "potentially handicapped" persons was three times as great, and similarly increased with age.

Table 3 - Results of Screening for Deafness, by Age and Sex of Persons Screened.

Screening result	TOTAL 1/	MALES, BY AGE IN YEARS					FEMALES, BY AGE IN YRS				
		2/ Total	<30	30-44	45-64	65+	2/ Total	<30	30-44	45-64	65+
Number screened	2016	920	173	403	286	56	1093	265	432	313	72
Normal	1485	600	150	303	137	8	883	252	394	209	21
Potentially handicapped	399	248	23	75	119	31	150	10	28	77	32
Handicapped	132	72	0	25	30	17	60	3	10	27	19
Deafness previously unknown	399	246	20	87	109	30	152	8	28	81	31
Potentially handicapped	323	206	20	66	96	24	116	6	21	65	23
Handicapped	76	40	0	21	13	6	36	2	7	18	8
Rate per 1000 persons screened											
Number screened	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Normal	737	652	867	752	479	143	808	951	912	668	292
Potentially handicapped	198	270	133	186	416	554	137	38	65	246	444
Handicapped	65	78	0	62	105	304	55	11	23	86	264
Deafness previously unknown	198	267	116	216	381	536	139	30	65	259	431
Potentially handicapped	160	224	116	164	336	429	106	23	49	201	320
Handicapped	38	43	0	52	46	107	33	7	16	58	111

1/ Includes unknown age and sex. 2/ Includes unknown age

Table 4 - Results of Screening for Deafness, by Race and Sex of Persons Screened.

Screening result	Total 1/ persons screened	WHITE			NON WHITE		
		Both sexes	Male	Female	Both sexes	Male	Female
Total screened	2016	1564	732	832	432	177	255
Rate per 1000 persons screened							
All results	1000	1000	1000	1000	1000	1000	1000
Normal	737	721	632	798	806	751	843
Potentially handicapped	198	209	293	144	148	198	114
Handicapped	65	70	85	58	46	51	43

1/ Includes unknown race and sex.

Sex. There were usually more "handicapped" men in each age group than women. The "potentially handicapped" persons, into which most of those with presbycusis belong, had a higher prevalence in men of all age groups. Of those men 45 years and over, more than 40 per cent were in this category.

Race. The non white groups had a lower proportion of "handicapped" and "potentially handicapped" persons than did the white groups.

Deafness previously known. 42 per cent of the "potentially handicapped" and 20 per cent of the "handicapped" persons were aware of their condition. In general, the older screenees were more aware of their deafness than the younger subjects.

Summary

A screening test for deafness will result in the discovery of a considerable number of previously unknown cases. However, special facilities seem necessary in order that the positive screenees may follow-up on the screening result.

REFERENCES

Best,C.H. and Taylor,N.B. (1955). The Physiological Basis of Medical Practice. The Williams and Wilkins Co. 6th Ed.

Houssay,B.A. et al. (1955). Human Physiology. McGraw-Hill Book Co.,Inc. 2nd Ed.

Sunderman,F.W. and Boerner,F. (1950). Normal Values in Clinical Medicine. W.B.Saunders Co.

U.S.Public Health Service, Washington, D.C. National Health Survey: Hearing Study Series, Bulletin No.5. *

*Not read in the original.

SCREENING FOR DENTAL CONDITIONS

The Baltimore clinic included a five minute dental examination. This procedure had been standardised so that different dentists would have comparable findings.

The primary purpose of this examination was to provide information on the prevalence of dental and related conditions in the population. The screenee was informed of the findings, however, and was likely to seek dental treatment when a correctible abnormality was present.

This procedure cannot be regarded as a typical screening test, since it must be carried out by professional personnel. Moreover, since dental conditions are highly prevalent, the procedure does not sort out a small proportion of persons needing treatment, but really directs the majority of persons attending the screening clinic to go on to their dentist for advice and treatment.

The procedure is probably effective in persuading many persons to seek treatment earlier than usual. However, it does not seem to be an economical way to use professional personnel. Dental hygienists may possibly be trained to carry out such examinations. If so, they are likely to detect at an earlier stage a considerable number of dental conditions, often of a progressive nature. The success of the procedure will again depend on the hygienists' ability to persuade screenees to seek correction of the defects. Further information is needed to determine the value of including a dental examination in a multiple screening procedure.

DISCUSSION AND CONCLUSIONS

Section C
Chapter 10

DISCUSSION AND CONCLUSIONS

This study has discovered a number of problems in multiple screening, some of which have not been previously mentioned in the published literature. There is now sufficient information to answer some of these problems. However, this study has inadequacies which leave unanswered many important questions on multiple screening.

The Baltimore clinic results give some idea of the yield of previously unknown confirmed cases which will result from the use of various tests. However, the figures available are based on the diagnostic findings and replies of the personal physicians of the referrals. These replies may have had a considerable degree of inaccuracy. This defect could have been corrected by referring all positive screenees to a facility which ensured adequate diagnostic examination. Since future programs in the United States will send positive screenees to general practitioners, however, it was felt that this study should use the same method of referral.

A further defect of the figures representing the yield of the tests is the lack of information on the proportion of confirmed cases likely to benefit by their early discovery. The figures may suggest, for instance, that a test carried out on screenees forty-five years and over will result in a considerable number of confirmed new cases. It is possible, however, that the cases discovered in this age group receive little help from treatment, and that the much smaller number of cases discovered by screening a younger group obtain more benefit

by their early discovery. Tests which seem adequate at present may later be shown to be worthless because the confirmed cases receive no benefit by their discovery.

Screening Tests Worthy of Future Use

The results of the Baltimore clinic and of other screening programs suggest that a number of tests are of value in future screening programs. These tests are:

Blood sugar tests for diabetes mellitus, preferably after the recent ingestion of glucose or food. The results of this procedure in the Baltimore clinic were unsatisfactory, and the reasons for this have been discussed. However, there is now considerable evidence that blood sugar tests give a satisfactory yield with a smaller proportion of false positives than do urine sugar tests.

Blood pressure measurement for hypertension, using higher screening levels for older age groups. The yield of previously unknown cases is quite high, although it is doubtful that many cases benefit by their early discovery.

Chest x-ray for heart disease is a useful screening test for persons 45 years of age and over. The yield below that age is quite low.

Six-lead electrocardiogram for heart disease. As used in Baltimore, this test is the most sensitive of the screening procedures for heart conditions. The results obtained in Baltimore were better than might be expected from the findings of other studies. The procedure is superior to the lead 1 tracing alone, but is of value mainly in those persons aged 45 years and over.

Haemoglobin level for anaemia, preferably using a fairly accurate photoelectric colourimetric method. The Baltimore and other studies suggest that the yield is mainly confined to women. However, the yield in men could be increased by raising the screening level to 13 or 14 gm. per cent.

Chest x-ray for tuberculosis. While the yield of previously unknown cases is relatively low, the importance and infectivity of the disease make this procedure worth while.

Screening Tests of Doubtful or No Value

Height and weight for obesity. While this test yields many previously unknown cases, the number who successfully lose weight and maintain the reduced weight is likely to be small.

Urine albumin for proteinuria. This test results in a large proportion of positive screenees who are not in need of treatment. The early detection of persistent proteinuria is of doubtful value.

Urine sugar. This test usually produces a considerable proportion of false positives, and is unnecessary when a blood sugar test is carried out.

Serological test for syphilis. This test has a very low yield of previously unknown cases in the general population of the United States. The test results in the referral of a considerable proportion of persons who do not require additional treatment.

Visual acuity test. This test produces a large yield of previously unknown cases. However, since early discovery neither delays the progress of nor cures the condition detected, there is some doubt as to the value of this procedure.

Hearing Acuity test. This test also yields many previously unknown cases. However, special diagnostic and treatment facilities may be necessary to ensure adequate follow-up of the positive screenees.

Multiple Screening and the American General Practitioner

Perhaps the greatest administrative difficulty in multiple screening in the United States, especially when the health department originates the program, is the maintenance of good relations with practicing physicians. American general practitioners view with some apprehension the entry of health departments into the field of early detection of non-infectious conditions.

Many strong and, at times, emotionally charged criticisms have been published in American medical journals. So far, the American Medical Association has taken no official stand on this subject. Moreover, the Association has prepared and distributed a booklet (American Medical Association, 1955) giving information on screening surveys which have been held up to the present time. This booklet has helped a number of local medical societies in the United States to play an active part in the design and administration of screening programs.

Most screening organizations take much trouble to consult with and get the active cooperation of practicing physicians. Nevertheless, there have apparently been difficulties, often glossed over or unmentioned in published articles, which might have a detrimental effect on other programs of the health department.

Public health administrators should therefore move slowly in suggesting such programs in the United States. In the future, if multiple screening is shown to be an effective and economical method of case finding, it may even be wise for the official public health administrator to take a back seat, allowing voluntary organizations or the private physicians themselves to administer such programs.

Other Administrative Problems

A large part of this thesis has been devoted to developing accurate and specific screening tests, which is one of the main administrative problems. When rescreening is necessary, a number of positive screenees will fail to take the rescreening test unless it can be administered at the time of the first visit. In the Baltimore clinic, the rescreening was done successfully in the case of the hearing and blood sugar tests. No solution was found, however, to permit the ECG and chest x-ray rescreening tests to be carried out during the first visit.

Methods must be developed to encourage positive screenees to go on for diagnostic examination, and to provide adequate medical facilities for the follow-up and treatment of these screenees. In the Baltimore clinic, a final interview with a public health nurse on the completion of the screening tests seemed to have some success in encouraging screenees to follow-up on their positive results.

Encouraging the attendance of those groups which most need screening is also a large administrative problem. A number of screening tests are suitable mainly for the older age groups,

who seem to be less interested in attending screening programs.

Reducing the cost of multiple screening seems essential before this procedure is suitable for large scale use. The Baltimore clinic does not provide useful cost figures because of its atypical nature, and no such figures are given in this thesis. However, excluding the research costs, about 57 shillings were spent in testing each screenee. Even if the clinic began continuous operation, it is unlikely that the cost could be reduced by more than 50 per cent. This cost would still be a great burden to any community which wished to initiate a large scale program.

Conclusion

Multiple screening is a sufficiently promising procedure to deserve much further development and study. However, multiple screening has not sufficiently developed to be suitable for application to large groups of the population. Programs which are held in the near future should concentrate mainly on clearing up the many problems which have not been satisfactorily solved. Foremost among these problems is a need for adequate diagnostic examinations to be performed on the negative as well as the positive screenees, in order to determine the proportion of missed cases. The other vast problem is the need to show whether or not positive screenees benefit by their early detection and diagnosis.

REFERENCE

American Medical Association (1955). A Study of Multiple Screening. Council on Medical Service, American Medical Association, 535 North Dearborn Street, Chicago 10, Illinois.