THE WIDESPREAD POPULARITY of using the ST-segment response to exercise to detect individuals with ischemic heart disease is based largely on epidemiologic data and arteriographic correlate studies. Epidemiologic studies of asymptomatic individuals have shown that the presence of a positive ST-segment response to exercise distinguishes a group of individuals at relatively high risk of developing coronary events over the next several years.1-5 Conversely, the presence of a negative test response distinguishes those individuals with a very low probability of such events. Arteriographic studies, in which the results of the exercise test response have been correlated with evidence of coronary luminal narrowing, also have shown that in symptomatic subjects the presence of an abnormal ST-segment response to stress in most instances correctly denotes the presence of important anatomic abnormalities of the coronary arteries.6-10

From this evidence, it has been concluded that the ST-segment response to exercise is highly accurate in predicting the presence or absence of coronary disease. However, this conclusion has been questioned in the light of recent coronary arteriographic studies of asymptomatic subjects with positive exercise tests. In such an asymptomatic population the predictive accuracy of a positive ST segment response (that is, the percentage of positive test results which correctly predict important angiographically-demonstrable coronary arterial disease [table 1]) is relatively low. Thus, in a population of asymptomatic subjects, Froelicher found that only 44% of all positive tests are true positives.11 We also found that only 37% of positive tests in asymptomatic subjects accurately reflect important coronary disease.12

It would seem that the results of these epidemiologic and arteriographic studies contradict one another. However, closer scrutiny of published data indicates that the disparity is only apparent. Thus, the low predictive accuracy of a positive exercise test in the angiographic studies of asymptomatic subjects is entirely compatible with the conclusions of epidemiologic studies, once the difference between risk ratio and predictive accuracy is appreciated (table 1). The results of epidemiologic studies8-14 (table 2) show that a group of asymptomatic subjects with positive test results is indeed at higher risk (10-15 times) of eventually developing coronary events than a group of asymptomatic subjects with negative tests (i.e., the risk ratio is high). This, however, should not be taken to imply that the large majority of subjects with positive tests have disease. In fact, the risk of subsequent coronary events in an asymptomatic individual with a positive test response (i.e. predictive accuracy) is low and comparable to the values found in the angiographic studies quoted above.11,12 Thus, predictive accuracy in the epidemiologic studies of groups of subjects followed from 3 to 6.3 years varies from 13.6% to 46% and averages, for the total number of 135 subjects, 25% (table 2).

The demonstration that such a large percentage of false positive exercise tests occurs among asymptomatic individuals, both in epidemiologic and in coronary angiographic studies, contrasts sharply with the excellent predictive accuracy of a positive test response in angiographic studies of symptomatic subjects8-10 and suggests that the predictive accuracy of a positive test may be dependent largely upon the population under study. This empirically derived hypothesis can be proven mathematically and is applicable to any diagnostic test. Thus it can be shown18 that the lower the prevalence of disease in a study population, the greater the likelihood that a positive test will be falsely positive, i.e., that the predictive accuracy of the test will be low. Conversely, in a population with high prevalence of disease, predictive accuracy will be high.

This is illustrated by the following example (table 3). Let us assume that both the specificity and sensitivity of a diagnostic test are 95%. If 1,000 patients are tested from a population in which the disease prevalence is 90% (as might exist in a population of patients referred to an institution because of symptoms of coronary disease) then 900 would have coronary disease and 100 would be free of disease. Since the test is 95% sensitive, then 855 of the 900 diseased patients will have a positive test. In addition, if specificity is 95%, five of the 100 patients without disease will have a falsely positive test result. Thus there would be 860 patients with a positive test, of whom 855 would actually have coronary disease, yielding a predictive accuracy of 99%.

In contrast, if 1,000 patients are tested from a population with a disease prevalence of only 2% (as might exist in an otherwise unsel ected group of asymptomatic individuals) then 20 would have coronary disease and 980 would be free of disease. Since the test is again 95% sensitive and 95% specific, 19 of the 20 subjects with disease will have a positive test and 49 (5%) of the 980 subjects without disease will have a falsely positive test result. Thus of 68 subjects with a positive test, only 19 will be true positives and the predictive
accuracy will be only 28%. (Reduction in sensitivity of the test to 75%, a figure commonly reported in both epidemiologic and diagnostic studies, reduces the predictive accuracy from 28% to 23%). The influence of disease prevalence on predictive accuracy of a positive test is graphically illustrated in figure 1.

Prevalence of coronary artery disease in published epidemiologic studies of groups of asymptomatic subjects in fact averages 4.4% (table 2). Conversely the prevalence of coronary disease is very high in groups of patients referred for evaluation of chest pain. The above theoretical considerations therefore offer a reasonable explanation for the apparently divergent results observed between the studies of symptomatic and asymptomatic subjects. Because the reliability of a positive exercise test result is largely dependent upon prevalence of disease in the study population, the high predictive accuracy in most published studies reflects the high prevalence of coronary disease in those groups of patients. In contrast, the epidemiologic and diagnostic studies reporting a low predictive accuracy for a positive stress test are based on results derived from groups of asymptomatic subjects, i.e., subjects with a low disease prevalence.

It is possible that some of the "false positive" responses occurring in asymptomatic patients may truly reflect myocardial ischemia, a physiologic abnormality, which may be present in the absence of angiographically demonstrable coronary artery disease; moreover, angiographic techniques may not detect existing large vessel coronary artery disease in all cases. In such situations, a positive electrocardiographic response to exercise might be the only objective manifestation of myocardial ischemia. However, the percentage of such patients in a population of subjects with positive exercise tests probably is quite low since epidemiologic studies also indicate that the chance is small that an overt coronary event will occur within five years in an asymptomatic subject with a positive ST-segment response to exercise.

This low predictive accuracy of positive test results in asymptomatic subjects might be improved by using more stringent criteria for selecting positive tests. For example, patients with ≥2 mm ST-segment depression occurring during exercise might be more likely to reveal important angiographic abnormalities than patients with <2 mm ST-segment depression. However, it can be shown mathematically that the predictive accuracy of any test remains

---

**Table 1. Glossary of Terms**

<table>
<thead>
<tr>
<th>Sensitivity — Percent of all patients with disease who manifest an abnormal test.</th>
<th>Specificity — Percent of negative results in subjects without the disease.</th>
<th>Predictive Accuracy — Percent of positive results that are true positive.</th>
<th>Risk Ratio — Percent of subjects with a positive test who manifest CAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>= [\frac{TP}{TP + FN} \times 100]</td>
<td>= [\frac{TN}{TN + FP} \times 100]</td>
<td>= [\frac{TP}{TP + FP} \times 100]</td>
<td>Percent of subjects with a positive test who manifest CAD</td>
</tr>
</tbody>
</table>

**Table 2. Incidence of Subsequent Development of Clinically Apparent Coronary Disease in an Asymptomatic Population**

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Pos. Test</th>
<th>% → CAD</th>
<th>Follow-up</th>
<th>Protocol</th>
<th>Criteria</th>
<th>Sensitivity</th>
<th>Predictive Accuracy</th>
<th>Specificity</th>
<th>Risk ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruce²</td>
<td>221</td>
<td>22</td>
<td>2.3%</td>
<td>5 yr</td>
<td>Maximal</td>
<td>Treadmill</td>
<td>≥1.0 mm</td>
<td>60%</td>
<td>13.6%</td>
<td>91%</td>
</tr>
<tr>
<td>1969</td>
<td>M</td>
<td>(34–67 yr)</td>
<td>(10%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aronow³</td>
<td>100</td>
<td>13</td>
<td>9%</td>
<td>5 yr</td>
<td>Maximal</td>
<td>Treadmill</td>
<td>≥1.0 mm</td>
<td>66%</td>
<td>46%</td>
<td>92%</td>
</tr>
<tr>
<td>1975</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(98 M, 2 F)</td>
</tr>
<tr>
<td>Froelicher⁴</td>
<td>451</td>
<td>39</td>
<td>3.8%</td>
<td>6.3 yr</td>
<td>Maximal</td>
<td>Treadmill</td>
<td>≥1.0 mm</td>
<td>58.8%</td>
<td>25.6%</td>
<td>93.6%</td>
</tr>
<tr>
<td>1974</td>
<td>M</td>
<td>(40–54 yr)</td>
<td>(8.6%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumming⁵</td>
<td>510</td>
<td>61</td>
<td>5.1%</td>
<td>3 yr</td>
<td>Maximal</td>
<td>Bicycle</td>
<td>≥2.0 mm</td>
<td>58%</td>
<td>25%</td>
<td>90.7%</td>
</tr>
<tr>
<td>1975</td>
<td>M</td>
<td>(40–65 yr)</td>
<td>(12%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** N = number; M = male; F = female; % → CAD = percent of total patient population subsequently manifesting coronary disease.

---

**Table 3. Relation of Predictive Accuracy to Disease Prevalence**

<table>
<thead>
<tr>
<th>I. 1) Disease prevalence = 90%</th>
<th>II. 2) Test 95% Sensitive and 95% Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>Number with Positive Test</td>
</tr>
<tr>
<td>900 diseased</td>
<td>855</td>
</tr>
<tr>
<td>100 nondiseased</td>
<td>5</td>
</tr>
<tr>
<td>1,000</td>
<td>860</td>
</tr>
</tbody>
</table>

**Abbreviations:** TP = true positive; TN = true negative; FP = false positive; FN = false negative; CAD = coronary artery disease.
restricted whenever the test is not 100% specific and is applied to a population with low disease prevalence. Thus if one assumes, using these more stringent diagnostic criteria, that specificity increases to as much as 99% and sensitivity is unchanged, predictive accuracy in a population with a coronary disease prevalence of 2% will increase to only 66%. However, test sensitivity would be compromised by the more stringent criteria and many subjects with significant disease would remain undetected. The decrease in sensitivity necessarily results in a reduction in predictive accuracy (fig. 1). That is, if sensitivity is only 50% instead of 95%, predictive accuracy falls from 66% to 50% in a population with disease prevalence of 2%. Indeed, recent epidemiologic studies by Froelicher and by Cumming suggest that when positive tests are defined by ≤2 mm ST-segment depression, predictive accuracy is even lower than these calculations suggest; predictive accuracy was found to be only 30% by Froelicher, and 25% by Cumming. Moreover, the reduction in predictive accuracy as a function of diminished sensitivity is more marked as disease prevalence diminishes (fig. 1). If the more stringent criterion of ≥2 mm ST-segment depression is adopted for use in symptomatic patients in an attempt to distinguish high risk patients (those with left main coronary artery disease or its equivalent), sensitivity also is compromised: about 50% of patients with such disease have either a normal ST-segment response to exercise or develop <2 mm depression. Of course, it is possible, though as yet unproven, that computer analysis of the ST-segment response to exercise will provide important improvement in both sensitivity and predictive accuracy.

Predictive accuracy of a positive stress test in asymptomatic subjects might be enhanced if a positive test was defined as one in which a previously negative exercise response converted to a positive response on serial testing. Such a possibility was suggested by the study of Doyle and Kinch, which analyzed the natural history of over 2000 subjects without overt evidence of coronary artery disease and with negative stress tests at entry to the study. It was shown that the subsequent development of a positive test delineated a group of subjects at very high risk of developing coronary events, namely angina pectoris, myocardial infarction, or sudden death. Thus, of 75 men who developed an abnormal test on serial testing, 85% developed overt coronary disease within the next five years. This contrasted to a 25% probability of developing coronary events in the group of subjects that had an abnormal test at entry into the study. Although other explanations are possible and confirmatory data are not available, these results suggest that the predictive accuracy associated with conversion from a negative to a positive response to exercise may be considerably higher than the predictive accuracy of a single positive exercise. Obviously, however, this hypothesis must be tested by additional prospective studies.

It should be emphasized that these concerns regarding the value of the ST-segment response to exercise in a symptomatic patients should not be interpreted as a brief against the use of exercise testing for other purposes. Thus, exercise testing is evidently of value for the determination of exercise tolerance, for the evaluation of symptoms and the effect of medical or surgical therapy and for the evaluation of arrhythmias.

In conclusion, we believe that the value of the ST-segment response to exercise stress testing as an aid in the diagnosis of coronary artery disease is considerably more circumscribed than is commonly believed. First, it is of limited diagnostic value in symptomatic patients with a history of typical angina pectoris or a previously documented myocardial infarction since, despite the high predictive accuracy of a positive test result in such patients, the diagnosis is virtually established by history alone. In addition, a negative test result in such patients would be of dubious value because of the relatively low sensitivity of the test in a symptomatic disease.
population. A possible exception to these limitations would be the presence of a strongly positive stress test (e.g., ≥ 2.0 mm ST depression) which might indicate left main coronary artery disease or its equivalent. However, as we have already pointed out, the test is insensitive when such criteria are adopted. Thus, detection of high risk subgroups of patients can only reliably be carried out by means of coronary arteriography.

The ST-segment response to exercise may have greater potential in patients with atypical chest pain syndromes. For example, if patients with obvious pain of pleuritic or chest wall origin are excluded, the prevalence of coronary disease in this group of patients may be as high as 30 to 50%. If this is true, figure 1 indicates that the predictive accuracy of a positive test, assuming a specificity of 95%, would be greater than 90%. However, there is no reason to suspect that the sensitivity of the test is any greater than that observed in patients with a more typical history of coronary disease; hence, a negative test result again would be of little help in ruling out the presence of coronary disease.

Finally, and most importantly, the test is of very limited diagnostic value when applied to asymptomatic subjects because of the large percentage of false positive results that occurs in this population. Thus a positive result in an asymptomatic patient presents the physician (and patient) with a dilemma. It is possible that informing the patient of a positive test result and of its implications might lead to better adherence to reasonable therapeutic programs aimed at correcting such risk factors as smoking, hypertension, or elevated serum cholesterol. However, the ethics involved in the use of such a threat are questionable when the odds are against the patient having significant coronary disease. Another consideration in such a situation is the serious psychological impact that a positive exercise test can have on a patient. Thus, many patients are unable to cope with the possibility of coronary disease and premature death, a particularly unfortunate situation considering the high probability that no disease is present. Even the use of more rigid criteria for a positive test result does not appear to result in a significant increase in the predictive accuracy of the test.

Thus, in order to distinguish between true positive and false positive ST-segment responses to exercise, a physician who elects to conduct routine screening exercise tests also should be prepared to proceed with coronary arteriography despite its attendant inconvenience, cost, and risk. It may be reasonable to adopt this latter alternative for those subjects whose occupation is such that an acute coronary event would place other individuals at risk. However, the appropriate management of the asymptomatic patient with coronary disease is still unclear, and at present, there is no evidence to indicate that operative intervention enhances survival in these patients. Hence, the premise that coronary arteriography should be performed in all asymptomatic subjects with positive exercise responses is highly questionable, since in the large majority of instances the findings will not importantly influence therapeutic decisions. In the end, it may be that many coronary angiograms will be performed not because prophylactic surgery is being considered, but to allay the anxiety the patient develops following his being informed that he has a positive exercise test and may have coronary disease. Given these considerations, it may be asked whether the decision to adopt exercise stress testing as a routine screening procedure is a wise one. On the other hand, though questionable as a routine procedure, such a policy might be reasonable to pursue in those subjects whose occupations are such that an acute coronary event might place other individuals at risk.

It is hoped that the concepts presented in this discussion will help to define the value of the ST-segment response to exercise in more modest, but perhaps more realistic, terms than have been accepted previously. Only when the strengths and weaknesses of the test are fully appreciated can it be properly employed or rejected in the assessment and management of the individual patient.

DAVID R. REDWOOD, M.D.
JEFFREY S. BORER, M.D.
STEPHEN E. EPSSTEIN, M.D.

References
Whither the ST segment during exercise.
D R Redwood, J S Borer and S E Epstein

Circulation. 1976;54:703-706
doi: 10.1161/01.CIR.54.5.703

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/54/5/703.citation