Incremental Value of the Exercise Test for Diagnosing the Presence or Absence of Coronary Artery Disease

Lee Goldman, M.D., M.P.H., E. Francis Cook, M.S., Nancy Mitchell, R.N., Margaret Flatley, B.S., Herbert Sherman, D.E.E., Robert Rosati, M.D., Frank Harrell, Ph.D., Kerry Lee, Ph.D., and Peter F. Cohn, M.D.

SUMMARY To determine the incremental value of the exercise test (ETT) for diagnosing coronary artery disease (CAD), we derived a multivariate logistic regression model for the pre-ETT prediction of CAD using data from 3840 patients at Duke University. We then applied the model to 324 patients at the Brigham and Women’s Hospital. Using seven clinical factors, the multivariate model had an 84% overall predictive accuracy on both the training (Duke) and the validation (Brigham) sets of patients. Three ETT factors (ST-segment change in patients not taking digitalis, absence of ST-segment change in patients taking digitalis, ETT stopped because of ECG or blood pressure changes) had incremental, significant predictive power, but overall predictive accuracy based on both clinical and ETT factors improved only to 87%. When the ETT result was important enough to move the probability of CAD across a potential therapeutic threshold, the direction of the change in probability was correct only two-thirds of the time. Thus, the ETT was of limited value in predicting the presence or absence of CAD after other easily obtainable clinical data were taken into account.

Although a general correlation between the results of exercise ECGs and angiographically defined coronary artery disease has been recognized for some time,1-3 many investigators have questioned the value of the exercise test in predicting the presence of coronary artery disease in individual patients.4-7 Recent studies have stressed the incremental value of the exercise ECG as it influences the probability of coronary artery disease,5-7 and attempts have been made to measure the information content of the exercise test itself.8 However, in recent studies,5-8 only one aspect of the exercise test, the amount of ST-segment depression, has been analyzed, despite evidence9-11 that several aspects of the exercise test might have independent predictive value. Furthermore, analyses that emphasized the magnitude of a change in probability often did not discriminate between changes in probability that would move a patient across a decision-making threshold and changes in probability that might confirm a diagnosis without moving the probability across such a threshold. Using the threshold approach to clinical decision-making, a test would have potential value when it might produce a result that could move a patient across a therapeutic threshold and would have actual management value when its result truly moved a patient across a threshold.12

With these principles in mind, we designed this study with several goals. First, we, like others,13-15 attempted to assign patients an accurate probability of coronary artery disease based on a multivariate analysis of historical, physical examination, and laboratory data that are readily available before exercise testing. Second, we measured the incremental value of information derived from the exercise test after such pretest factors had been taken into account. Third, we determined how often patients could potentially move across a therapeutic threshold based on the results of the exercise test, how often patients actually moved across such a threshold based on their own test results, and how often such changes were accurate.

Methods

The Duke University Medical Center Coronary Artery Disease Data Bank and the Harvard University–Brigham Coronary Artery Disease Data Bank have been described in detail.16, 17 Both data banks contain over 900 items of clinical, angiographic, hemodynamic, surgical and follow-up data from each patient entered. At the Duke University Medical Center, patients were eligible for this project if they underwent their initial cardiac catheterization between July 1969 and August 1980 and did not have valvular heart disease or emergency surgery. In a few cases, all pertinent variables were not recorded before the exercise test; thus, 3840 patients are included in this report. At the Brigham and Women’s Hospital, similar criteria have been in effect since 1980. During the prior 3 years, Brigham patients were entered if they were patients of the full-time cardiology faculty, if they were not being catheterized as part of a hospital admission that included time spent in the coronary care unit, and if they had coronary disease at catheterization. Using these criteria, 842 patients were routinely entered between January 1977 and May 1981; 250 of these performed an exercise test within 6 months before or several weeks after the catheterization and did not have an interval infarction and thus were included in this
report. Because patients who were found not to have coronary disease at catheterization at Brigham and Women's Hospital in 1977–1979 had originally been excluded, the Harvard-Brigham Data Bank was supplemented with all 79 patients who were catheterized during the same period and who otherwise met our original 1977–1979 criteria except that they did not have coronary artery disease at catheterization. The Brigham and Women's Hospital cohort included 44 patients (14%) who were taking digitalis and 187 (58%) who were taking propranolol. The patients' average age was 53 years, and 249 patients (77%) were male.

At Brigham and Women's Hospital, patients were exercised using the standard Bruce protocol, with each stage lasting 3 minutes. Patients were encouraged to exercise to their maximal tolerance, but tests were terminated if necessary because of symptoms, ECG changes (marked ST-segment changes, arrhythmias, new bundle branch block or atrioventricular block), or blood pressure changes (hypotension or marked hypertension). Only 12 patients had normal resting and exercise ECGs and exercised to less than 85% of their maximal predicted heart rate. Standard 12-lead ECGs were taken in the supine and upright positions and after hyperventilation before exercise, every 3 minutes during exercise, and again at 0, 1, 3, 5 and 8 minutes after exercise or until the ECG returned to baseline.

At both centers, coronary arteriograms of the left and right coronary systems were taken in multiple views. Coronary artery disease was defined by the presence of a 70% or greater occlusion in at least one coronary vessel. Angiograms were always interpreted by two or more experienced angiographers; readers from the two centers have met together in the past to assure that the interpretations would be as uniform as possible at both sites. Among the Harvard-Brigham patients, 79 (24%) had no significant coronary lesions, 33 (10%) had one-vessel disease, 79 (24%) had two-vessel disease, and 133 (41%) had three-vessel disease. Twenty-seven patients (8%) had significant left main lesions, and 153 (47%) had abnormal left ventricular contraction.

Statistical Methods

To develop pre-exercise-test probabilities of coronary artery disease, the Duke University Medical Center Data Bank was used to derive a simple model for predicting the results of coronary arteriography. Historical, physical examination and pre-exercise-test laboratory data were screened for testing in the analysis. The historical variables included age; sex; history of arrhythmia; typical, atypical or nonanginal chest pain; severity of congestive heart failure and angina (New York Association functional classes I to IV); duration (in years) of symptoms of coronary disease; number of chest pain episodes per week; presence of nocturnal chest pain; precipitation of chest pain by effort; family history of coronary artery disease; clinical history of myocardial infarction; response of chest pain to vaso-dilators; presence of preinfarction angina or progressive chest pain; history of smoking, obesity, hyperlipidemia, cerebrovascular or peripheral vascular disease, menopause, or hypertension. The physical examination variables included systolic blood pressure, diastolic blood pressure, lung rales, local chest tenderness, atrial gallop ($S_1$), ventricular gallop ($S_2$), presence of systolic ejection murmur or of other systolic murmurs, presence of peripheral bruits, absence of peripheral pulses, and xanthoma. The ECG variables included diagnostic Q waves compatible with old infarction, QRS abnormalities, ST-T changes on the resting ECG, intraventricular conduction disturbances, left-axis deviation, right-axis deviation, right bundle branch block, left ventricular hypertrophy, rhythm other than normal sinus, and premature ventricular contractions. Patients with left bundle branch block on the resting ECG were excluded. Finally, the laboratory variables considered were cholesterol level and cardiomegaly on the chest radiograph.

A logistic regression model for predicting the probability ($P$) of coronary artery disease (i.e., the presence of one or more major coronary vessels with a 70% or greater obstruction) was constructed in a stepwise manner, using variables listed above as candidates for the model. The form of this model was

$$\log \frac{P}{1-P} = B_0 + B_1X_1 + \ldots + B_nX_n$$

where $X_1, \ldots, X_n$ represented the chosen variables, $B_0$ represented a constant, and $B_1, \ldots, B_n$ represented the relative weights assigned to the chosen variables. At each step of the logistic regression procedure, the variable that was most significantly associated with the presence of coronary artery disease was selected, conditioning on all variables previously chosen. All variables that had significant incremental correlations with the presence of coronary disease and that improved the overall accuracy of the model were included.

The logistic regression model that was based on the Duke University Medical Center patients was then applied to the Harvard-Brigham Coronary Artery Disease Data Bank patients to ascertain its performance on an independent set of patients. The model’s high degree of accuracy on the Harvard-Brigham patients was especially noteworthy because it was based on patients from a different institution.

Having developed a simple, accurate model for predicting the pre-exercise-test probability for coronary artery disease, we then attempted to determine the incremental predictive information that could be obtained from an exercise test. We examined the following exercise test variables to determine if they could add significantly to the model that was based on the preexercise-test factors: the greatest ST-segment change measured in any lead on any of the exercise or postexercise ECGs; duration of exercise in seconds; maximal heart rate during exercise; maximal blood pressure during exercise; maximal double product (heart rate $\times$ blood pressure); treadmill terminated because of angina; treadmill terminated because of
nonanginal chest pain; treadmill terminated because of ECG changes, arrhythmia, atrioventricular or bundle branch block, marked hypertension or hypotension; treadmill terminated because of shortness of breath or dizziness; treadmill terminated because of fatigue, nausea, leg pain or arm pain; angina during stress; treadmill complications. We also constructed numerous combination and interaction variables, including combinations of maximal ST change with the use of digitalis, with duration of exercise, and with the maximal heart rate achieved. Factors that were not routinely recorded and hence were not considered in the analysis included changes in R-wave amplitude, treadmill time at onset of ST depression, sloping ST depression, and duration of ST depression after exercise. Only three patients had new ST-segment elevation during the exercise test; all of these were in leads that showed resting ST elevation, and two of the patients had exercise-induced ST depression in other leads as well. The incremental ST elevation in the one patient without concomitant ST depression was coded as ST change, as were equivalent amounts of ST depression.

All exercise test variables were then examined to determine their incremental ability to predict coronary disease after the factors included in the pre-exercise-test model had been taken into account. This examination was accomplished by treating all exercise test variables as candidates for a second stepwise logistic regression procedure that was forced to contain the pre-exercise-test score (i.e., the linear combination of the pre-exercise-test variables and their weight derived from the Duke patients). Actually, this restriction proved unnecessary because the pre-exercise model was more significantly associated with the presence of coronary disease than was any single exercise test variable. All exercise test variables that had incremental significant correlations with the presence of coronary disease and that improved the overall predictive accuracy of the logistic regression model were included.

Finally, the impact of exercise test information for the patients in the Harvard-Brigham Coronary Artery Data Bank was evaluated by comparing the pre- and postexercise probabilities of coronary artery disease in reference to various theoretical decision-making thresholds. Using the actual observed range of the exercise test variables that added significant incremental predictive information in the Harvard-Brigham patients, the number of patients “at risk” of crossing a threshold was calculated. Then, the number of patients who actually crossed a threshold based on their own exercise test results, as well as the number who crossed correctly, were determined. The overall accuracy of our predictive models was calculated using the Q statistic, which measures how well a probability estimate agrees with the presence or absence of a finding. The goodness of fit for the multivariate models was also measured using the Goodman-Kruskal gamma coefficient, which assesses the degree to which diseased and nondiseased patients are separated by their estimated probabilities. Both the Q statistic and the Goodman-Kruskal gamma coefficient were analogous to a correlation coefficient in that +1 indicates perfect prediction and separation respectively.

A supplemental logistic regression analysis was performed to determine what exercise test variables provided incremental information in the subset of Harvard-Brigham patients who did not have Q waves or a clinical history of a myocardial infarction. The results of this supplemental analysis were similar to the analysis on all patients except that the diminution in sample size reduced the significance of two of the exercise test factors to a borderline level, clearly by also introducing a substantial β error. Thus, all 324 Harvard-Brigham patients were included in our logistic regression analysis, and we used our threshold analysis to compare the value of the exercise test in patients with and without Q waves or a clinical history of a myocardial infarction.

Results

Of the 3840 eligible patients in the Duke University Medical Center Data Bank, 2493 patients (65%) had significant coronary artery disease and 1347 patients (35%) did not. Multivariate logistic regression analysis showed that seven pre-exercise-test factors in these patients had statistically significant univariate and multivariate correlations with the presence of coronary artery disease at catheterization and contributed to an improvement in overall predictive accuracy (table 1). Six percent of patients with supposedly diagnostic Q waves and 8% of patients with a clinical history of infarction did not have significant coronary disease at catheterization. Factors that had strong univariate correlations with the presence of coronary disease but did not add incremental predictive power included diabetes, positive family history for coronary artery disease, obesity, hypertension, presence of peripheral vascular disease, resting ST-T changes, and effort-related pain. Using a cutoff point whereby patients with multivariate estimated probabilities of greater than 50% would be predicted to have coronary disease and patients with probabilities below 50% would be predicted not to have coronary disease, the seven-factor logistic regression equation had a sensitivity of 88%, a specificity of 76%, and an overall accuracy of 84% when applied retrospectively to the same 3840 patients from whose data it was derived.

Prospective Application of the Pre-exercise-test Model to the Harvard-Brigham Patients

All seven significant factors identified in the Duke University patients had remarkably similar sensitivities and specificities in the Harvard-Brigham patients (table 1). The seven-factor multivariate logistic regression equation using the weights derived from the Duke patients yielded very similar results in the Harvard-Brigham patients. The Duke-based model actually had a slightly higher sensitivity (93%) and an identical overall accuracy (84%) for predicting the presence of coronary disease in Harvard-Brigham patients; how-
ever, its specificity of 57% was lower in the Harvard-Brigham patients than in the Duke patients ($p < 0.01$).

**Incremental Value of the Exercise Test**

After the Duke-based model had been applied to the Harvard-Brigham patients, only three of the many available exercise test factors contributed incremental multivariate significance: the maximal ST change in patients who were not taking digitalis, the absence of ST change in patients who were taking digitalis, and an exercise test that was stopped because of ECG or blood pressure changes (table 2). Other exercise test variables considered individually or in combination could not add statistically significant incremental value to the resulting logistic regression model. Despite the incremental statistical significance of the three exercise test variables to the multivariate model, the additional clinical value for classifying Harvard-Brigham patients above or below a 50% probability of coronary artery disease was minimal (table 3), and overall accuracy increased only from 84% to 87%.

If all potential pre-exercise-test and exercise test variables were considered at once, the single best predictor of the presence or absence of coronary disease was the ST change on the exercise test. However, ST change correlated with the typicality of the patient’s pain ($r = 0.26$), with the patient’s age ($r = 0.21$), and with the patient’s score on the seven-factor model ($r = 0.19$; all $p < 0.001$). Thus, the exercise test result was not independent of other clinical factors, but was correlated with them to the extent that much of the exercise test’s predictive value was duplicated by factors that were less expensive to gather.

**Threshold Analysis**

Of the 324 patients in the Brigham series, 184 (57%) had pre-exercise-test probabilities of coronary artery disease that ranged between 0.01 and 0.92 and hence could have had exercise test results that could either be so positive (for example, 6 mm of ST-segment depression) or so negative that the probability of coronary artery disease might be moved across a 50% threshold based on the exercise test result. Using a similar analysis for other thresholds, we calculated the minimal and maximal pre-exercise-test probabilities for which an exercise test could move the probability of disease across a given threshold (table 4). However, strongly positive exercise tests did not often occur in patients with low pre-exercise-test probabilities of coronary artery disease. Thus, many patients who might have had exercise test results that would move them across a particular probability threshold in fact often did not. In our series, none of the patients with pre-exercise-test probabilities below 25% had exercise test results that moved them across a 50% threshold (table 5). Among

**Table 1. Significant Independent Predictors of Coronary Artery Disease**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Univariate correlation with CAD in Duke patients</th>
<th>Univariate correlation with CAD in Harvard-Brigham patients</th>
<th>Assigned weight in pre-exercise-test multivariate model*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Is chest pain typical for angina pectoris?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0. No</td>
<td>848 (34%)</td>
<td>1113 (83%)</td>
<td>88 (36%)</td>
</tr>
<tr>
<td>1. Yes</td>
<td>1645 (66%)</td>
<td>234 (17%)</td>
<td>157 (64%)</td>
</tr>
<tr>
<td>B. Does ECG have Q waves thought to be diagnostic of an old MI?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0. No</td>
<td>1499 (60%)</td>
<td>1283 (95%)</td>
<td>135 (55%)</td>
</tr>
<tr>
<td>1. Yes</td>
<td>994 (40%)</td>
<td>64 (5%)</td>
<td>110 (45%)</td>
</tr>
<tr>
<td>C. Patient’s sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0. Female</td>
<td>402 (16%)</td>
<td>693 (51%)</td>
<td>39 (16%)</td>
</tr>
<tr>
<td>1. Male</td>
<td>2091 (84%)</td>
<td>654 (49%)</td>
<td>206 (84%)</td>
</tr>
<tr>
<td>D. Does patient have a clinical history of an acute MI?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0. No</td>
<td>1271 (51%)</td>
<td>1245 (92%)</td>
<td>102 (42%)</td>
</tr>
<tr>
<td>1. Yes</td>
<td>1222 (49%)</td>
<td>102 (8%)</td>
<td>143 (58%)</td>
</tr>
<tr>
<td>E. Has the patient smoked at least ½ pack per day in the past 5 years?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0. No</td>
<td>669 (27%)</td>
<td>679 (50%)</td>
<td>93 (38%)</td>
</tr>
<tr>
<td>1. Yes</td>
<td>1824 (73%)</td>
<td>668 (50%)</td>
<td>152 (62%)</td>
</tr>
<tr>
<td>F. Cholesterol level (mg%; mean ± sd)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>238 ± 50</td>
<td>223 ± 45</td>
<td>240 ± 46</td>
<td>213 ± 51</td>
</tr>
<tr>
<td>G. Age (years; mean ± sd)</td>
<td>51 ± 9</td>
<td>47 ± 9</td>
<td>54 ± 9</td>
</tr>
</tbody>
</table>

*For factors A–E, a patient was assigned the variable weight if the factor had the value labeled 1. For cholesterol level and age, the actual value was multiplied by the weight. The resulting logistic regression equation was thus: \[
\log \left( \frac{P}{1-P} \right) = \Sigma F_A W_A + \Sigma F_G W_G - 7.6234,
\]

where $F_A$ . . . $F_G$ represent factors and $W_A$ . . . $W_G$ represent their assigned weights. Abbreviations: CAD = coronary artery disease; MI = myocardial infarction.
TABLE 2. Independently Significant Factors from the Exercise Tolerance Test for Predicting Coronary Artery Disease*  

<table>
<thead>
<tr>
<th>Univariate correlation in Brigham patients</th>
<th>Multivariate weight</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAD (n = 245)</strong></td>
<td><strong>No CAD (n = 79)</strong></td>
</tr>
<tr>
<td><strong>ST change in patients not taking digitalis</strong></td>
<td></td>
</tr>
<tr>
<td>0. 0 mm or taking digitalis</td>
<td>77 (31%)</td>
</tr>
<tr>
<td>1. 1 mm</td>
<td>54 (22%)</td>
</tr>
<tr>
<td>2. 2 mm</td>
<td>53 (22%)</td>
</tr>
<tr>
<td>3. 3 mm</td>
<td>41 (17%)</td>
</tr>
<tr>
<td>4. ≥ 4 mm</td>
<td>20 (8%)</td>
</tr>
<tr>
<td><strong>Absence of ST change in a patient taking digitalis</strong></td>
<td></td>
</tr>
<tr>
<td>0. Not taking digitalis or ST change in a patient taking digitalis</td>
<td>240 (98%)</td>
</tr>
<tr>
<td><strong>ETT stopped because of ECG or blood pressure changes</strong></td>
<td></td>
</tr>
<tr>
<td>0. No</td>
<td>171 (70%)</td>
</tr>
<tr>
<td>1. Yes</td>
<td>74 (30%)</td>
</tr>
</tbody>
</table>

*Where post-ETT probability of CAD = P, and log ( \( \frac{P}{1-P} \) ) = 0.84 (pre-ETT score) + \( F_A W_A + F_B W_B + F_C W_C \) - 0.57, and \( F_A \ldots F_C \) represent factors and \( W_A \ldots W_C \) represent their assigned weights.

Abbreviations: CAD = coronary artery disease; ETT = exercise tolerance test.

patients with pre-exercise-test probabilities of 25–49%, 10 patients (29%) had exercise test results that moved them across a 50% threshold; three patients (9%) had exercise tests that were sufficiently negative to reduce the probability to below 10%, and only two patients (6%) had exercise tests that were sufficiently positive to move the probability above greater than 90%. Among the 151 patients with pre-exercise-test probabilities above 90%, none had exercise test results that moved them across a 50% threshold, and only one patient had an exercise test result that moved the probability across a 75% threshold probability of having coronary disease.

Thus, the number of patients who crossed individual threshold probabilities after the exercise test varied among thresholds (table 6). For the entire 324 patients, the number crossing a threshold ranged from 3–16%. Unfortunately, an average of 36% of the patients who crossed a threshold actually moved in a direction that incorrectly predicted whether or not they had coronary artery disease.

When the threshold analysis was limited to patients without Q waves on the ECG and no clinical history of myocardial infarction, the percentage of patients crossing a threshold was 5–19%, depending on the thresholds, and an average of 75% of the patients who crossed a threshold did so in the correct direction (table 6). The main value of the exercise test in patients without a Q wave or a clinical history of infarction was to move patients with pre-exercise-test probabilities of 50–90% across a 90% threshold probability of having coronary disease.

**Discussion**

Numerous studies have provided data on the factors that correlate with the risk that a previously healthy person will develop either a myocardial infarction or clinically diagnosed angina pectoris.22–25 Consistently important predictive factors have included serum cholesterol level, blood pressure, age, cigarette smoking, glucose tolerance, and electrocardiographic abnormalities. In studies that included females, gender was an independently significant risk factor.22 Other factors, such as obesity,23 parental history of coronary disease,25 and psychosocial characteristics,24,25 have also been reported to be important in some multivariate prospective studies. Our study population was different, however, in that we analyzed patients who had chest pain syndromes and who subsequently underwent an exercise test and cardiac catheterization. Other investigators who have studied patient populations

TABLE 3. Impact of the Exercise Tolerance Test for Improving the Overall Classification of Brigham Patients

<table>
<thead>
<tr>
<th>All Brigham patients (n = 324)</th>
<th>Brigham patients without Q waves or clinical history of MI (n = 149)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-ETT</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>93%</td>
</tr>
<tr>
<td>Specificity</td>
<td>57%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>84%</td>
</tr>
<tr>
<td>Accuracy coefficient</td>
<td>0.45</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td>0.72</td>
</tr>
</tbody>
</table>

(Goodman-Kruskal gamma)

Abbreviations: MI = myocardial infarction; ETT = exercise tolerance test.
similar to ours have reported conflicting data regarding the ability to predict the presence of significant coronary artery disease based on clinical factors. Salel and colleagues, in a study of patients referred for catheterization to evaluate chest pain, dyspnea or heart murmur, found higher coronary risk profile scores based on the factors and weights from the Framingham study in patients who had coronary artery disease than in those without coronary disease. However, Salel and colleagues did not report the precise discriminating ability of the coronary risk profile, and based on their reported means and standard errors it appears that there may have been substantial overlap between the two groups. Vlietstra and colleagues also reported that age, sex, cigarette smoking and cholesterol levels were significant discriminators between patients with coronary artery disease at angiography vs those with chest pain syndromes but no coronary disease; a family history of coronary disease and the presence of hypertension or diabetes were less helpful in their study, much as they were in ours. Based on the raw data provided in their article, approximately 81% of patients were correctly classified by the overall multivariate discriminant function. In our series, four classic cardiac risk factors (sex, age, cholesterol level and smoking), one electrocardiographic factor (the presence of a Q-wave), and two factors related to symptoms of coronary disease (clinical history of myocardial infarction and typical angina) correctly classified 84% of patients in the group from which the multivariate function was derived and in a prospective set of patients at a second hospital. The slightly better performance of our multivariate model was probably related to the inclusion of the three factors not used by Vlietstra et al. (electrocardiographic Q waves, a clinical history of a myocardial infarction, and the typicality of anginal pain), although patients with prior myocardial infarctions were included in their study. Dimsdale and colleagues used clinical and epidemiologic variables similar to ours and had an overall 90% accuracy for the retrospective classification of the same 100 patients from whose data the discriminant function score was derived. Cohn and colleagues used seven factors to predict whether patients with symptoms of chest pain would have coronary disease at angiography. However, their multivariate model, which had an overall accuracy of 91% in an independent testing sample of 100 patients, required several noninvasive tests.

### Incremental Value of an Exercise Tolerance Test

The values and limitations of exercise testing for the diagnosis of coronary artery disease have been discussed, but any analysis of the incremental value of exercise stress testing will vary, depending on how the exercise test is considered. For example, the Bayesian analysis of Weiner and colleagues considered tests to be either normal or abnormal, without indicating the degree of abnormality. The low impact that they reported for the exercise test might have been a function of an inadequate consideration of the degree of positivity. In their Bayesian analyses, Diamond and Forrester and Rifkin and Hood considered the degree of positivity on the exercise test based on ST-segment depression, but neither considered other exercise test variables that might contribute additional information. Conversely, most of the studies that have shown that many other factors in the exercise test correlated with the presence and severity of coronary disease did not consider the pre-exercise-test probability of coronary disease based on clinical data.

In their comprehensive study of the exercise test, Fisher and colleagues examined clinical and exercise test variables to see which factors were important in men with definite angina, probable angina, or nonspecific chest pain. Of the clinical factors, age was import-

### TABLE 4. Potential Value of the Exercise Tolerance Test for Moving Patients Across Therapeutic Thresholds

<table>
<thead>
<tr>
<th>Threshold probability</th>
<th>Range of pre-ETT probabilities that patient could be moved across threshold</th>
<th>No. of the 324 patients who could possibly cross threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.0007 to 0.46</td>
<td>56 (17%)</td>
</tr>
<tr>
<td>0.25</td>
<td>0.003 to 0.76</td>
<td>106 (33%)</td>
</tr>
<tr>
<td>0.50</td>
<td>0.01 to 0.92</td>
<td>184 (57%)</td>
</tr>
<tr>
<td>0.75</td>
<td>0.04 to 0.98</td>
<td>252 (78%)</td>
</tr>
<tr>
<td>0.90</td>
<td>0.12 to 0.994</td>
<td>284 (88%)</td>
</tr>
</tbody>
</table>

Abbreviation: ETT = exercise tolerance test.

### TABLE 5. Actual Value of the Exercise Tolerance Test for Moving Patients Across Therapeutic Thresholds

<table>
<thead>
<tr>
<th>Threshold probability</th>
<th>Percent of patients whose ETT actually moved them across the threshold*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-ETT &lt; 0.10</td>
</tr>
<tr>
<td>0.10</td>
<td>27%</td>
</tr>
<tr>
<td>0.25</td>
<td>9%</td>
</tr>
<tr>
<td>0.50</td>
<td>0%</td>
</tr>
<tr>
<td>0.75</td>
<td>0%</td>
</tr>
<tr>
<td>0.90</td>
<td>0%</td>
</tr>
</tbody>
</table>

*These figures are all-inclusive. Thus, of the patients with pre-ETT probabilities of coronary disease that were below 10%, 27% had ETT results that moved their probabilities above 10%, including 9% of patients who moved to above 25%.

Abbreviation: ETT = exercise tolerance test.
EXERCISE TEST DIAGNOSIS OF CAD/Goldman et al.

Table 6. Net Yield of Exercise Tolerance Test Results for Various Therapeutic Thresholds

<table>
<thead>
<tr>
<th>Threshold probability</th>
<th>No. of patients moving across threshold based on ETT result</th>
<th>No. of patients correctly moving across threshold based on ETT result</th>
<th>Net correct yield because of ETT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients without Q waves or clinical history of MI</td>
<td>Patients without Q waves or clinical history of MI</td>
<td>Patients without Q waves or clinical history of MI</td>
</tr>
<tr>
<td></td>
<td>All patients (n = 324)</td>
<td>All patients (n = 148)</td>
<td>All patients</td>
</tr>
<tr>
<td>0.10</td>
<td>8</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>0.25</td>
<td>14</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>0.50</td>
<td>23</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>0.75</td>
<td>43</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>0.90</td>
<td>53</td>
<td>28</td>
<td>33</td>
</tr>
</tbody>
</table>

Abbreviations: ETT = exercise tolerance test; MI = myocardial infarction.

Important in all three groups of men, cigarette smoking was important only in men with nonspecific chest pain, and the cholesterol level was not an independent predictor in any group. Fisher et al. concluded that exercise testing provides more diagnostic information than clinical data alone in the patients with definite angina and probable angina, but probability-curve analysis showed that the amount of incremental information provided by the exercise test was modest. Similarly, we found that the exercise test was the best single piece of information for predicting whether patients had coronary disease, but its incremental value in classifying patients after the clinical factors had been considered was limited. Even if our analysis was applied only to patients without prior myocardial infarction, as did Fisher et al., the relative incremental value of the exercise test remained limited.

The limited incremental value of the exercise test was further underscored by our method. We required that our seven-factor, Duke-based, pre-exercise-test model be validated on the Harvard-Brigham patients, but we reported the value of the exercise test factors based solely on their utility in the Harvard-Brigham patients from whose data the factors were derived. Because a predictive model commonly performs better on the patients from whose data it was derived than on other sets of patients, the incremental value of the exercise test factors might be overestimated by our method.

Threshold Analysis

As discussed by Pauker and Kassirer, the impact of diagnostic tests can best be understood using the threshold approach to clinical decision making. Using such an approach, treatment would be given and the test would not be ordered if the probability of disease was already above a therapeutic threshold and if it would be impossible for the test to move a patient across such a threshold. Thus, a test would be performed only when it might move a patient across a therapeutic threshold. We have expanded this principle by including not only the possibility that a test might move a patient across a therapeutic threshold, but also the probability that such a test result might be found in an individual patient. This modification allows a physician to weight the potential benefits of a test on a continuous scale rather than having to consider it as a yes-no phenomenon.

There is no agreement among clinicians as to the appropriate probability threshold for management decisions in coronary disease. For example, if the probability of coronary disease is 25%, should a symptomatic patient still be given a medical trial and then be catheterized if symptoms persist? Although clinicians and patients may respond differently to the same probability estimate, the explicit delineation of probabilities should be valuable for physician-patient dialogue and for future research into patient management.

Using our probability approach, the incremental value of an exercise test for the diagnosis of coronary artery disease appears to be limited. When all patients were considered, only 7% of patients moved across a 50% threshold probability; because six of these patients moved in the wrong direction, only 3.5% of patients (i.e., one of 28 patients tested) had a new, correct classification based on the exercise test result. When the analysis was limited to patients with no clinical history of acute myocardial infarction and without Q waves, 11% of patients moved across the 50% probability threshold, but a net of only 5% of patients (i.e., one of 20 patients tested) had new, correct classifications. Our more detailed analysis showed that many patients who might cross a threshold did not. Thus, it would be unusual for an exercise test to move a patient from below a 25% probability of coronary disease to above a 50% probability in actual practice. Patients with pre-exercise probabilities of 25–75% had the greatest absolute changes in probability based on the test result, but commonly the disease was neither established nor excluded.
The basic threshold principles could be applied to the sequential use of several tests and also to the use of an individual test for other purposes such as the prediction of three-vessel disease or left main coronary artery stenosis. Because clinicians often order exercise tests for more than one purpose, it would also be possible to sum the potential value of an exercise test for different purposes and different thresholds. Although we did not study the value of exercise tests for objectively determining exercise capacity, inducing arrhythmias, predicting multivessel disease, or estimating prognosis, our data suggest that exercise tests appear to be of limited value for predicting the presence or absence of coronary disease in an individual patient in whom routine historical and laboratory data are available.

References

Appendix
Calculation of Pre- and Post-exercise-test Probabilities of Coronary Disease

Example 1
A 55-year-old man has typical angina, no history of myocardial infarction, and no Q waves on his ECG. He smokes 1 pack of cigarettes...
per day and has a cholesterol of 320 mg%. His pre-exercise-test probability \( P_1 \) of coronary disease is

\[
\log \frac{P_1}{1-P_1} = 1.54 + 1.79 + 0.75 + 2.88 + 3.69 - 7.62 = 3.03
\]

\[
P_1 = \frac{e^{3.03}}{1 + e^{3.03}} = 0.95
\]

If he has 2-mm ST depression during exercise and stops because of symptoms, his post-exercise-test probability \( P_2 \) of coronary disease is:

\[
\log \frac{P_2}{1-P_2} = 0.84 (3.03) + 0.96 - 0.57 = 2.94
\]

\[
P_2 = \frac{e^{2.94}}{1 + e^{2.94}} = 0.95
\]

Thus, the exercise test has not changed the probability of coronary disease at all.

**Example 2**

Consider the same patient as in example 1, except the chest pain is atypical angina. Then the pre-exercise-test probability \( P_1 \) would be:

\[
\log \frac{P_1}{1-P_1} = 1.54 + 0.75 + 2.88 + 3.69 - 7.62 = 1.24
\]

\[
P_1 = \frac{e^{1.24}}{1 + e^{1.24}} = 0.78
\]

Again assuming the same exercise test result as in example 1, the post-exercise-test probability \( P_2 \) would be:

\[
\log \frac{P_2}{1-P_2} = 0.84 (1.24) + 0.96 - 0.57 = 1.43
\]

\[
P_2 = \frac{e^{1.43}}{1 + e^{1.43}} = 0.81
\]

Thus, even in this patient with several risk factors and atypical angina, the exercise test would make only a very small change in the probability of coronary disease.

---

**The Independent Value of Exercise Thallium Scintigraphy to Physicians**

**MARK HLATKY, M.D., ELIAS BOTVINICK, M.D., AND BRUCE BRUNDAZE, M.D.**

**SUMMARY**

To determine the effect of exercise myocardial scintigraphy with thallium-201 on diagnostic accuracy and the need for coronary angiography, consecutive patients with a variety of clinical presentations were identified. Clinical summaries, including a detailed history, physical examination, and complete data from a standard treadmill exercise test, were presented to 91 cardiologists. The cardiologists assessed the probability of coronary disease and the need for coronary angiography. They were then presented the results of thallium scintigraphy and revised their assessments if warranted.

Scintigraphy significantly increased the cardiologists' diagnostic accuracy beyond that attained with other clinical information \( (p < 0.0001) \). The change in accuracy varied from + 4% to + 20% in different patient groups, and was greatest in patients with atypical angina and a positive exercise ECG. Ratings of the need for coronary angiography changed from −13% to +21% in different patient groups. We conclude that exercise thallium scintigraphy can provide independent diagnostic information and influence the need for coronary angiography.

**STRESS** perfusion scintigraphy with thallium-201 is a new diagnostic test for coronary artery disease. Scintigraphy documents a different manifestation of myocardial ischemia than the exercise ECG and may improve diagnostic accuracy.

The use of exercise testing for the diagnosis of coronary disease is controversial. There is considerable disagreement as to whether exercise testing adds new and significant information to that already available from clinical examination. Documentation of the independent value of thallium scintigraphy is therefore particularly important, as this procedure supplements, rather than replaces, the exercise ECG. Moreover, scintigraphy exposes the patient to a small amount of radiation, requires 1–5 hours to perform, and costs several hundred dollars.

The present study was designed to measure how much information thallium perfusion scintigraphy adds to that already available from the history, physical examination and standard exercise testing. The study also examines the effect of perfusion scintigraphy on the need for coronary angiography, an important aspect of clinical management.
Incremental value of the exercise test for diagnosing the presence or absence of coronary artery disease.
L Goldman, E F Cook, N Mitchell, M Flatley, H Sherman, R Rosati, F Harrell, K Lee and P F Cohn

Circulation. 1982;66:945-953
doi: 10.1161/01.CIR.66.5.945

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1982 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

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/66/5/945

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/