Simple Treadmill Score To Diagnose Coronary Disease*

Vinod Raxwal, MD; Katerina Shetler, MD; Anthony Morise, MD; Dat Do, MD; Jonathan Myers, PhD; J. Edwin Atwood, MD; and Victor F. Froelicher, MD

Objective: **Our aim was to derive and validate a simplified treadmill score for predicting the probability of angiographically confirmed coronary artery disease (CAD).**
Background: **The American College of Cardiology/American Heart Association guidelines for exercise testing recommend the use of multivariable equations to enhance the diagnostic characteristics of the standard treadmill test. Most of these equations use complicated statistical techniques to provide diagnostic estimates of CAD. Simplified scores derived from such equations that require physicians only to add points have been developed for pretest estimates of disease and for prognosis. However, no simplified score has been developed specifically for the diagnosis of CAD using exercise test results.**
Methods: **Consecutive patients referred for evaluation of chest pain who underwent standard treadmill testing followed by coronary angiography were studied. A logistic regression model was used to predict clinically significant (≥ 50% stenosis) CAD and then the variables and coefficients were used to derive a simplified score. The simplified score was calculated as follows: (6 × maximal heart rate code) + (5 × ST-segment depression code) + (4 × age code) + angina pectoris code + hypercholesterolemia code + diabetes code + treadmill angina index code. The simplified score had a range from 6 to 95, with < 40 designated as low probability, between 40 and 60 was intermediate probability, and > 60 was high probability for CAD.**
Results: **A total of 1,282 male patients without a prior myocardial infarction underwent exercise treadmill testing and coronary angiography in the derivation group, and there were 476 male patients in the validation group from another institution. The area under the receiver operating characteristic curve (± SE) for the ST-segment response alone was 0.67 as compared to 0.79 ± 0.01 for the diagnostic score (p > 0.001). The prevalence of significant disease for the men was 27% in the low-probability group, 62% in the intermediate-probability group, and 92% in the high-probability group, which was similar to the prevalence in the validation group, with 22%, 58%, and 92% in low-, intermediate-, and high-probability groups, respectively. The low-probability group had < 4% prevalence of severe disease. In both populations, 7 more patients out of 100 were correctly classified than with the use of ST-segment criteria. When used as a clinical management strategy, the score has a sensitivity of 88% and a specificity of 96%.**
Conclusion: **This simplified exercise score that estimates the probability of CAD can be easily applied without a calculator and is a useful and valid tool that can help physicians manage patients presenting with chest pain.** ***(CHEST 2001; 119:1933–1940)***

Key words: clinical prediction rules; coronary artery disease; exercise testing; scores

Abbreviations: AUC = area under curve; CAD = coronary artery disease; MET = metabolic equivalent; ROC = receiver operating characteristic; VA = Veterans Affairs

*From the Divisions of Cardiovascular Medicine, Stanford University Medical Center, and the Veterans Affairs Palo Alto Health Care System (Drs. Raxwal, Shetler, Do, Myers, Atwood, and Froelicher), Palo Alto, CA; and University of West Virginia (Dr. Morise), Morgantown, WV.
Manuscript received December 20, 2000; accepted December 22, 2000.
Correspondence to: Victor Froelicher, MD, Cardiology Division (111C), Veterans Affairs Palo Alto Health Care System, 3801 Miranda Ave, Palo Alto, CA 94304; e-mail: vicmd@aol.com

The exercise ECG test is the recommended test for diagnosing coronary artery disease (CAD) in patients with chest pain and at intermediate probability for CAD.[1] Statistical techniques that combine the patient's medical history, symptoms of chest pain, hemodynamic data, and exercise ECG response have been demonstrated to better predict CAD than a single ECG criterion like ST-segment depression.[2] Studies have shown that the diagnostic value of exercise testing can be improved by considering several factors in the test interpretation,[3–7] but issues remain about their portability.[8] In addition, even though the American College of Cardiology/American Heart Association guidelines recommend that equations should be used to increase the value of the test,[9] many clinicians have not used them because of their complexity. Resolution of these two limitations to the application of equations (that is, portability and complexity) would be especially helpful today when more than half of the exercise tests are performed by noncardiologists. As health-care costs continue to increase, emphasis will grow on using the exercise tests as the gatekeeper to expensive interventions.[10] Furthermore, there is an awareness of the need to apply scores for better decision making.[11] Therefore, we have attempted to develop and demonstrate the portability of an easily applied clinical score that provides a management strategy for the noncardiologist evaluating patients with suspected coronary disease.

Materials and Methods

Patients

Veterans Affairs Score Development Population: Eight thousand male patients underwent treadmill testing at two Veterans Affairs (VA) medical centers between 1987 and 1998. Of these patients, 3,454 were evaluated for chest pain with coronary angiography within 3 months of treadmill testing. Patients with previous cardiac surgery or angiography, valvular heart disease, left bundle-branch block, paced rhythms, or Wolff-Parkinson-White syndrome on their resting ECG were excluded from the study. Since neither medications nor resting ST-segment depression of < 1 mm have been shown to alter the diagnostic accuracy of the test,[12] exclusions were not made for these reasons. Prior cardiac surgery was the predominant reason for exclusion of patients who underwent exercise treadmill testing during this time period. Patients with previous myocardial infarction by history or by Q-wave criteria were also excluded from the study in order to minimize the effects of measurement bias. A total of 1,282 patients met the inclusion and exclusion criteria of our study.

A complete clinical history was obtained at the time of exercise treadmill testing, and $> 90\%$ of patients were in the intermediate pretest probability level specified by the guidelines. While much of these data were gathered prospectively using computerized forms,[13,14] some of the patients initially studied had incomplete data requiring retrospective chart review.

West Virginia University Validation Population: Nine hundred eighteen consecutive patients (52% were male patients) with complete data underwent exercise treadmill testing at West Virginia University Hospital between 1981 and 1998 to evaluate chest pain possibly due to coronary disease. All patients underwent coronary angiography within 4 months of the exercise treadmill test, and the same exclusion criteria were applied as for the VA population. The 476 men from this group were utilized to validate the exercise test score derived from the VA patients. The 442 women were utilized to determine if an exercise score specific to women was needed. A complete clinical history was obtained prospectively at the time of exercise treadmill testing using the same computerized database.

Exercise Treadmill Testing

The 12-lead maximal effort exercise tests were performed utilizing standard graduated treadmill protocols consistent with American Heart Association guidelines.[1] Patients were encouraged to give a maximal effort but not to allow their angina to reach levels higher than previously experienced. The results were analyzed and reported utilizing a computerized database at all three institutions (EXTRA; Mosby Publishers; Chicago, IL).[15] The ST-segment response considered was the most horizontal or downsloping ST-segment depression in any lead except aVR during exercise or recovery. An abnormal response was defined as ≥ 1 mm of horizontal or downsloping ST-segment depression. No test result was classified as indeterminate,[16] treatment with medication was not withheld, and no maximal heart rate targets were applied.

Coronary Angiography

Coronary artery narrowing was visually estimated and expressed as percent lumen diameter stenosis. Patients with a 50% narrowing in one or more of the following were considered to have significant angiographically confirmed CAD: the left anterior descending, left circumflex, right coronary arteries or their major branches, or a 50% narrowing in the left main coronary artery. Severe disease was considered to include two vessels with this criterion if one is proximal left anterior descending or three vessels or left main. The 50% lesion criterion was chosen to be consistent with the cooperative trialists choice.[17]

Decisions for cardiac catheterization were consistent with clinical practice. Analyses were performed with the investigators blinded to clinical and angiographic results.

Statistical Methods

A statistical technique was used to separate subjects into those with and without significant angiographically confirmed disease based on clinical and measured exercise variables in the two derivation populations of 1,282 men and 442 women using Number Cruncher Statistical System (NCSS Statistical Software; Kaysville, UT). Forward selection was used with entry at a significance level > 0.05. The general linear logistic regression model used took the following form:

$$\text{probability (0 to 1)} = 1/(1 + e^{-(a+bx+cy\ldots)}),$$

where a is the intercept, b and c are coefficients, and x and y are variable values. Logistic regression is advantageous since dichotomous and continuous variables can be considered together and the output ranges from 0 to 1 representing the probability of disease being present. The equation fits a sigmoid curve that is a common biological relationship between a risk factor and disease.

How well the model separates patients with and without a given outcome (CAD) was assessed by means of the area under a receiver operating characteristic (ROC) curve, which ranges from 0 to 1, with 0.5 corresponding to no discrimination (*ie*, random performance) and 1.0 to perfect discrimination.

Score Derivation

While multivariable logistic regression techniques have much to recommend them, the equations they produce are complicated and it is difficult to understand the relative importance of selected variables. Moreover, the equations take the form of exponentials and require the use of a calculator in order to estimate the probability of disease. To decrease the complexity of these equations, it is possible to use the variables chosen in logistic regression in a simple linear score. We first coded all variables with the same number of intervals so that the coefficients would be proportional. Then we coded the bin with the larger value to be associated with higher probability of disease. For instance, if 5 is the chosen interval, dichotomous variables are 0 if not present and 5 if present and continuous variables like age and heart rate are coded in 5 bins by appropriate ranges. All codes then would be directly related to probability (*ie*, a heart rate code of 5 would be a low heart rate while an age code of 5 would be for the oldest individuals), and the smallest coefficient is associated with the least important variable. The multiplier of this least important variable was reduced to unity and the other coefficients into their proportional weight or importance by dividing each coefficient by the smallest coefficient. This makes the relative importance of the selected variables very obvious. Such techniques have been applied before for Cox hazard function equations. This approach results in a very simple linear score in which the health-care provider merely compiles the variables in the score, multiples by the appropriate number, and then adds up the products. Surprisingly, these simple linear scores have the same ROC areas as the more complicated equations requiring the calculation of exponentials.

Three steps were used to derive the new treadmill score. Initially, we tested the validity/portability of the pretest score of Morise et al[18] by comparing it to an equation derived in our population (ROC area under curve [AUC] = 0.71 vs 0.73, no significant difference).[18] Second, we derived a non-ECG equation by considering all of the hemodynamic variables, appropriate products, and their differences from baseline (*ie*, metabolic equivalents [METs], systolic BP, maximal heart rate, and treadmill angina index) in a logistic regression model (ROC AUC = 0.68). Third, we entered the Morise pretest score, the non-ECG equation, and amount of exercise-induced ST-segment depression into a logistic model. The resulting equation exhibited a ROC AUC of approximately 0.79. In order to further simplify, the variables previously chosen were reconsidered in a logistic model that eliminated some variables. This logistic equation was then used to create our final simplified score.

Change in systolic BP and METs were eliminated when considered with the pretest variables and exercise-induced ST-segment depression. Cigarette smoking, obesity, and family history were eliminated from the pretest variables originally present in the Morise pretest score. A simple linear score was derived by multiplying the coefficient of the variables from the multivariable equation with the variable code. The men's score was calculated as follows: ($6 \times$ maximal heart rate code) + ($5 \times$ ST-segment depression code) + ($4 \times$ age code) + angina pectoris code + hypercholesterolemia code + diabetes code + treadmill angina index code.

This diagnostic score did not perform well for the 442 symptomatic women (AUC < 0.65) and so a female-specific score was derived. This score requires validation in a large sample of women at another institution since the VA population is 98% male.

Coding of the variables is illustrated in Figure 1, which can be carried on an index card.

Results

Population Characteristics

Of the 1,282 veterans included in this study, 759 patients (59%) had clinically significant CAD, 302

Variable	*Circle response*	*Sum*
Maximal Heart Rate	Less than 100 bpm = 30	
	100 to 129 bpm = 24	
	130 to 159 bpm =18	
	160 to 189 bpm =12	
	190 to 220 bpm =6	
Exercise ST Depression	1-2mm =15	
	> 2mm =25	
Age	>55 yrs =20	
	40 to 55 yrs = 12	
Angina History	Definite/Typical = 5	
	Probable/atypical =3	
	Non-cardiac pain =1	
Hypercholesterolemia?	Yes=5	
Diabetes?	Yes=5	
Exercise test induced Angina	Occurred =3	
	Reason for stopping =5	
	Total Score:	

Males

Choose only one per group

<40=low prob

40-60= intermediate probability

>60=high probability

FIGURE 1. Coding of the variables. bpm = beats per minute; prob = probability.

patients (24%) had multivessel disease, and 523 patients (41%) were without any CAD; the 476 men in the validation group from West Virginia University Medical Center had a 46% prevalence of CAD (Table 1). Overall, the CAD group of patients was older compared to patients without CAD in both the populations. The CAD groups had a higher prevalence of hypertension, diabetes, and hypercholesterolemia. In the derivation group, 52% had hypertension, 42% had hypercholesterolemia, and 34% had typical angina pectoris. The validation sample patients were younger; 30% of them had hypertension and hypercholesterolemia. The prevalence of diabetes was similar in both groups. Smoking was more prevalent in the validation sample. Both groups were similarly medicated: about a third were receiving β-blockers, < 3% were receiving digoxin, almost 40% were receiving calcium antagonists, and > 50% were receiving nitrates.

Exercise Test Results

As shown in Table 2, from the exercise test, mean ST-segment deviation was 0.7 mm, the average MET was 7.5%, and 42% of the patients developed angina during the exercise. Abnormal treadmill test results were reported in 38% of the patients in the derivation group as compared to 32% in the validation group. Angina was the reason for stopping the test in 14% of the patients. Maximal systolic BP during exercise was similar in both the groups with and without CAD, with a mean of 169 mm Hg.

Scores

Since the Morise pretest score demonstrated its portability by showing similar results to our pretest equation, we decided to use it. The Morise et al[18] score had a range of 0 to 24 points and an ROC AUC of 0.72. Posttest scores were generated for all 1,282 patients in the derivation group and 476 patients in the validation group referred for treadmill test for the diagnosis of CAD. There was a steady increase in prevalence of CAD as the score increased, with a range of 6 to 95 points. We considered the method of Kotler and Diamond[19] for cut point determination and then chose to define three groups as having low, intermediate, or high probability. The Kotler and Diamond[19] method is based on the assumption that there are two threshold levels of disease probability: (1) a lower threshold level below which the number of false-positive responses exceeds the number of true-positive responses, and (2) an upper threshold level above which the number of false-negative responses exceeds the number of true-negative responses. In our derivation population, low probability was defined as a score of < 40, intermediate was between 40 and 60, and high probability was > 60. We also calculated the Duke prognostic treadmill score since it has been used as a diagnostic score.[20]

Table 3 demonstrates the probability of CAD in the three groups (*ie*, low, intermediate, and high). In the derivation group, 25% of the patients were classified as low probability, 55% as intermediate probability, and 20% as high probability. The prevalence of any CAD in the low-probability group was 27%, 62% in the intermediate-probability group, and 92% in the high-probability group, which was comparable to the validation group, with 22%, 58%, and 92% in low-, intermediate-, and high-probability groups, respectively. The low-probability groups had a < 4% prevalence of severe CAD. Thus, besides exhibiting portability, our score did not have a

Table 1—*Clinical Characteristics in Male Patients With and Without CAD in the Derivation Group and the Validation Group**

Variables	Derivation Group (VA) CAD (n = 759, 59%)	Derivation Group (VA) No CAD (n = 523)	Validation Group (West Virginia) CAD (n = 220, 46%)	Validation Group (West Virginia) No CAD (n = 256)
Age, yr	61.8 ± 8.8	55.1 ± 11.2	58 ± 10	48 ± 11
Body mass index, kg/m^2	28.0 ± 4.6	27.8 ± 4.8	28 ± 4	29 ± 6
Hypertension	55 (256)	49 (417)	34 (74)	26 (67)
Diabetes	17 (131)	11 (59)	21 (47)	10 (25)
Family history of CAD	42 (365)	45 (233)	49 (108)	48 (122)
Hypercholesterolemia	48 (365)	34 (177)	32 (70)	29 (75)
Abnormal treadmill test result	51 (386)	20 (105)	47 (103)	20 (50)
Smokers	30 (229)	36 (190)	39 (86)	45 (116)
Symptom status				
Typical angina	43 (328)	21 (109)	35 (76)	18 (45)
Atypical angina	44 (330)	64 (335)	38 (83)	41 (105)
Nonanginal chest pain	8 (60)	8 (40)	27 (61)	41 (106)

*Data are presented as mean ± SD or percentage (No.) of subjects.

Table 2—*Exercise Test Results**

Variables	Derivation Group CAD (n = 759, 59%)	Derivation Group No CAD (n = 523)	Validation Group CAD (n = 220, 46%)	Validation Group No CAD (n = 256)
Maximal heart rate, beats/min	125 ± 22	136 ± 24	139 ± 20	154 ± 21
Maximal systolic BP, mm Hg	169 ± 30	170 ± 27	166 ± 24	169 ± 22
Δ Systolic BP, mm Hg	39 ± 27	46.5 ± 23	35 ± 21	42 ± 18
Maximal double product, × 1,000	21 ± 6.2	23 ± 6.4	23 ± 5.4	26 ± 5.3
METs	6.8 ± 2.7	8.4 ± 3.4	6.9 ± 2.7	8.3 ± 2.9
Mean ST-segment depression, mm	1.0 ± 1.1	0.4 ± 0.7	1.1 ± 1.1	0.6 ± 0.9
Exercise angina score				
1 point (angina during test)	31 (233)	24 (125)	11 (24)	27 (68)
2 points (angina reason for stopping the test)	19 (144)	6 (34)	33 (72)	2 (5)

*Data are presented as mean ± SD or percentage (No.) of subjects.

calibration problem (that is, cut points had similar probabilities in different populations).

Areas Under the ROC Curve

For predicting CAD, Morise et al[18] had an ROC AUC of 0.71 for the pretest score, 0.69 for the non-ECG score, and 0.67 for exercise-induced ST-segment depression. The area under the ROC curve for both the posttest scores was greater (0.79). The area under the ROC curve in both the derivation and validation groups is comparable (Table 4). Scores were also generated for 442 symptomatic women referred for diagnosis of CAD at the University of West Virginia.

Sensitivity, Specificity, and Predictive Accuracy

In Table 5, we compare the sensitivity and specificity of the standard ST-segment analysis to cut points in the scores that match the specificity of ST-segment alone. Predictive accuracy allows calculation of how many more patients will be correctly classified by one method compared to another. For instance, 7 more patients out of 100 tested will be correctly classified using the score compared to ST-segment depression alone, with the score having a predictive accuracy of 69% as contrasted to 62% that of ST-segment depression alone. If we consider that the intermediate group will eventually be correctly classified by further testing (including cardiac catheterization), then the sensitivity and specificity of this management approach are 88% and 96%, respectively.

Discussion

Recommendations have been made for assigning patients to low-, intermediate-, or high-probability groups based on clinical criteria in order to provide a strategy for patient management.[21] The use of a score for diagnostic purposes represents a compromise between the simplicity of designating arbitrary high probability, intermediate and low probability, and the accuracy of detailed logistic regression models.[22] Probability of disease subgrouping is appealing because the classification scheme is easily remembered and assignment to such groups has implications for management; for instance, whether or not to refer to a subspecialist. With pressures being exerted on physicians to only refer patients for more expensive procedures when it is absolutely necessary, this approach gives them additional basis for making the correct decision. We have demonstrated that our scores function as well or better than expert cardiologists, and so they offer a valid second opinion to the nonspecialist.[23]

Table 3—*Prevalence of CAD in Each of the Probability (for Disease) Stratification Groups**

Probability	Derivation Group Total Population, No. (n = 1,282)	Derivation Group Prevalence Any CAD (n = 747)	Validation Group Total Population, No. (n = 476)	Validation Group Prevalence Any CAD (n = 230)
Low (score < 40)	325	88 (27.1)	203	45 (22.1)
Intermediate (score 40 to 60)	696	433 (62.2)	198	116 (58.5)
High (score > 60)	254	235 (92.5)	75	69 (92)

*Data are presented as No. (%) unless otherwise indicated.

Table 4—*AUC of the ROC Curves for the Treadmill Equations and Scores**

Variables	Derivation Group: Any CAD, Score	Derivation Group: Any CAD, Equation	Derivation Group: Severe CAD, Score	Derivation Group: Severe CAD, Equation	Validation Group: Any CAD	Validation Group: Equation
Diagnostic score	0.79	0.79	0.79	0.78	0.80	0.78
Duke treadmill score	0.73		0.75		0.72	0.78
Morise pretest score	0.71		0.68		0.73	0.75
Pretest VA	0.73	0.74	0.67	0.69		

*Severe CAD = two vessel if one is proximal LAD or three or more vessels with 50% or greater lesions or left main. Any CAD = any 50% or greater lesion in one or more vessel.

Specific recommendations are appropriate for respective probability groups with the caveat that physician judgment is still paramount in the decision process. Table 6 displays the paradigm for the clinical reaction to the estimated probability of CAD. There is no need for immediate further testing for the patients in the low-probability group, since less than 1 in 4 patients will have clinically significant CAD and less than 1 in 25 patients would have severe disease. The low-probability patient can be reassured that symptoms are most likely not due to CAD. However, if the symptoms do not abate, good clinical judgment should be utilized (*ie*, repeat testing perhaps with imaging). In addition, a prognostic score can be used to reassure the low-probability patient as well as the physician.[20] However, the patients assigned to the high-probability group may need an intervention if clinically appropriate and are potential candidates for coronary angiography. In the high-probability subgroup of patients, the use of antianginal treatment is indicated. In the group of patients with intermediate probability of CAD, there is need for the other tests, such as stress echocardiography or nuclear angiography to clarify diagnosis, and antianginal medications may be tried.

This study is seminal for several reasons. While prognostic and pretest scores have been adopted, there has been no prior diagnostic score developed from an appropriate logistic model. Secondly, portability has always been an issue for such equations and scores and we have cross-validated the Morise pretest score and the new male exercise test score in a distinctly different population from which they were derived.

Limitations

Since workup bias was not limited by protocol in this study, our results may be affected. However, we and others have anecdotally noted that because more and more patients are undergoing coronary angiography irrespective of the exercise treadmill test result, the importance of eliminating workup bias in this setting has lessened. It is encouraging to see that our results and population characteristics are similar to the only study to reduce workup bias by protocol.[24] We demonstrated that this score did not discriminate in women, but we have a specific new exercise score for women that we plan to validate.

Conclusion

This clinical scoring method is an accurate and simple method for categorizing patients with suspected coronary disease into clinically meaningful groups for which decisions concerning patient management can be based. We have demonstrated that the score is portable and is diagnostically superior to standard exercise testing interpretation.

Appendix

Morise Pretest Score

age code + (angina pectoris code $\times$ 5) + (diabetes $\times$ 2) + hypertension + smoking now + hypercholesterolemia + family history of CAD + obesity,

Table 5—*Sensitivity, Specificity, and Predictive Accuracy of Scores Compared to Exercise-Induced ST-Segment Depression (Cut Points Matched to Specificity of ST-Segment Depression)*

Variables	Cut Point Value	Sensitivity, %	Specificity, %	Predictive Accuracy, %
Exercise ST-segment depression	1 mm	50	80	62
Morise pretest score	13 points	50	75	60
Diagnostic score	50 points	61	80	69
Duke treadmill score	1 point	54	81	65

Table 6—*Paradigm for the Clinical Reaction to the Estimated Probability of any CAD*

Probability*	Clinical Reaction
Low	Patient is reassured that symptoms are most likely not due to CAD (less than 1 in 4 patients have clinically significant CAD and less than 1 in 25 patients have severe CAD); good clinical judgment leads to further testing if symptoms persist or return.
Intermediate	Requires other tests, such as a stress echocardiography or nuclear angiography to clarify diagnosis; antianginal medications tried.
High	Antianginal treatment indicated; intervention clinically appropriate; angiography may be required.

*$\geq$ 50% occlusion.

where age $< 40 = 3$ points, age between 40 years and 55 years = 6 points, and age > 55 years = 9 points. For estrogen status, 3 points were subtracted for positive status and 3 points were added for negative status. Typical chest pain = 5 points, atypical chest pain = 3 points, nonanginal chest pain = 1 point, and no chest pain = 0 points. For diabetes mellitus, 2 points were added and 1 point was added for each of the other five risk factors (hypertension, present smoking, hypercholesterolemia, family history of CAD, and obesity).

Multivariable Pretest Equation for Diagnosing Any CAD Derived in Our Population

$-$ 2.9 + (0.55 $\times$ age code) + (0.21 times] angina pectoris code) + (0.13 $\times$ hypercholesterolemia code) + T-wave abnormality code + diabetes code $-$ standing heart rate code

Multivariable Posttest Equation for Any CAD

$-$ 4.36 + (0.47 $\times$ depression code) + (0.56 $\times$ heart rate code) + (0.39 $\times$ age code) + (0.14 $\times$ angina pectoris code) + (0.14 $\times$ hypercholesterolemia code) + (0.12 $\times$ angina index code) + diabetes code,

where angina pectoris (definite/typical angina pectoris = 5 points, probable/atypical = 3 points, noncardiac pain = 1 point, none = 0 points); hypercholesterolemia (yes = 5 points, no = 0 points); ST-segment depression ($<$ 1 mm = 0 points, 1 to 2 mm = 3 points, $>$ 2 mm = 5 points); diabetes (yes = 5 points, no = 0 points); age ($<$ 40 years = 0 points, 40 to 55 years = 3 points, $>$ 55 years = 5 points); treadmill angina index (index of 0 = 0 points; index of 1 = 3 points, index of 2 = 5 points); maximal heart rate ($<$ 100 beats/min = 5 points, 100 to 130 beats/min = 4 points, 130 to 150 beats/min = 3 points, 160 to 190 beats/min = 2, 190 to 220 beats/min = 1 point, $>$ 220 beats/min = 0 points); METs ($<$ 3 = 5 points, 3 to 6 = 4 points, 6 to 9 = 3 points, 9 to 12 = 2 points, 12 to 15 = 1 point, $>$ 15 = 0 points); δ pressure, *ie*, change of BP from the baseline during exercise ($<$ 20 mm Hg = 5 points, $>$ 20 mm Hg = 0 points).

Duke Treadmill Score

Duration of exercise in minutes (5 $\times$ maximal ST-segment deviation during or after exercise in millimeters) $-$ (4 $\times$ the treadmill angina index), where angina index has a value of 0 if the patient had no angina during exercise, 1 if the patient had nonlimiting angina, and 2 if angina was the reason the patient stopped exercising.

Glossary

- Equation – mathematical representation of the result of a multivariable statistical technique that attempts to discriminate those with and without disease
- Code – a numerical value for the variables included in an equation or score
- Score – a simplified version of an equation that only requires adding or subtracting of coded points
- Multiple logistic model – a multivariable statistical technique that attempts to discriminate those with and without disease and provides a probability of being in the diseased group from 0 to 1 calculated by a log equation
- ROC – receiver operator characteristic curve is a graphic representation of the relationship between sensitivity and specificity for a diagnostic test
- AUC – area under the ROC curve is a measure of how well the model separates patients with and without a given outcome (CAD). The AUC ranges from 0 to 1, with 0.5 corresponding to no discrimination (*ie*, random performance), 1.0 to perfect discrimination, and values $<$ 0.5 to worse-than-random performance.
- Portability – ability of a score or equation to discriminate in other than the population in which it is derived
- Calibration – how well the cut points of a score or equation correlate with actual disease probabilities in different populations

References

1 Fletcher GF, Froelicher VF, Hartley LH, et al. Exercise standards: a statement for health professionals from the American Heart Association. Circulation 1990; 82:2286–2321. Revised Circulation 1995; 91:580–632

2 Yamada H, Do D, Morise A, et al. Review of studies utilizing multi-variable analysis of clinical and exercise test data to predict angiographic coronary artery disease. Prog Cardiovasc Dis 1997; 39:457–481

3 Kansal S, Roitman D, Bradley EL Jr, et al. Enhanced evaluation of treadmill tests by means of scoring based on multivariate analysis and its clinical application: a study of 608 patients. Am J Cardiol 1983; 52:1155–1160

4 Fisher LD, Kennedy JW, Chaitman BR, et al. Diagnostic quantification of CASS (Coronary Artery Surgery Study) clinical and exercise test results in determining presence and extent of coronary artery disease: a multivariate approach. Circulation 1981; 63:987–1000

5 Hollenberg M, Budge WR, Wisneski JA, et al. Treadmill score quantifies electrocardiographic response to exercise and improves test accuracy and reproducibility. Circulation 1980; Feb 61:276–285

6 Cohn K, Kamm B, Feteih N, et al. Use of treadmill score to quantify ischemic response and predict extent of coronary disease. Circulation 1979;59:286–296

7 Berman JL, Wynne J, Cohn PF. A multivariate approach for interpreting treadmill exercise tests in coronary artery disease. Circulation 1978; 58:505–512

8 Laupacis A, Sekar N, Stiell I. Clinical prediction rules: a review and suggested modifications of methodologic standards. JAMA 1997; 277:488–494

9 Gibbons RJ, Balady GJ, Beasely JW, et al. ACC/AHH guidelines for exercise testing. J Am Coll Cardiol 1997; 30:260–315

10 Marcus R, Lowe R, Froelicher VF, et al. The exercise test as

gatekeeper: limiting access or appropriately directing resources? Chest 1995; 107:1442–1446
11 Swets JA, Dawes RM, Monahan J. Better decisions through science. Sci Am 2000; 283:82–87
12 Fearon W, Lee D, Froelicher VF. The effect of resting ST segment depression on the diagnostic characteristics of the standard exercise test. J Am Coll Cardiol 2000; 35:1206–1211
13 Ustin J, Umann T, Froelicher V. Data management: a better approach. Physicians Computers 1994; 12:30–33
14 Froelicher V, Shiu P. Exercise test interpretation system. Physicians Computers 1996; 14:40–44
15 Shue P, Froelicher V. Extra: an expert system for exercise reporting. J Noninvasive Testing 1998; II-4:21–27
16 Reid M, Lachs M, Feinstein A. Use of methodologic standards in diagnostic test research. JAMA 1995; 274:645–651
17 Yusuf S, Zucker D, Peduzzi P, et al. Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomized trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. Lancet 1994; 344:563–570
18 Morise AP, Haddad J, Beckner D. Development and validation of a clinical score to estimate the probability of coronary artery disease in men and women presenting with suspected coronary artery disease. Am J Med 1997; 102:350–356
19 Kotler TS, Diamond GA, Exercise thallium –201 scintigraphy in the diagnosis and prognosis of coronary artery disease. Ann Intern Med 1990; 113:684–702
20 Shaw LJ, Peterson ED, Shaw LK, et al. Use of a prognostic treadmill score in identifying diagnostic coronary disease subgroups. Circulation 1998; 98:1622–1630
21 Braunwald E, Mark DB, Jones RH, et al. Unstable angina: diagnosis and management; clinical practice guideline No. 10 (amended). Rockville, MD: Agency for Health Care Policy and Research and the National Heart, Lung, and Blood Institute, Public Health Service, US Department of Health and Human Services. May 1994:14–22AHCPR; publication No. 94–0602
22 Diamond GA, Staniloff HM, Forrester JS, et al. Computer assisted diagnosis in the noninvasive evaluation of patients with suspected coronary artery disease. J Am Coll Cardiol 1983; 1:444–455
23 Lipinski M, Atwood E, Osterberg L, et al. Can physicians diagnose coronary disease as well as scores? AHA scientific sessions [abstract]. Circulation 1999; 100(suppl 1):505
24 Froelicher VF, Lehmann KG, Thomas R, et al. The electrocardiographic exercise test in a population with reduced workup bias: diagnostic performance, computerized interpretation, and multivariable prediction. Ann Intern Med 1998; 128(12 pt 1):965–974