
Treadmill Score Quantifies Electrocardiographic Response to Exercise and Improves Test Accuracy and Reproducibility

MILTON HOLLENBERG, M.D., W. ROGER BUDGE, M.D., JUDITH A. WISNESKI, M.D., AND EDWARD W. GERTZ, M.D.

SUMMARY We have developed a sensitive, accurate and highly reproducible treadmill exercise score (TES) that grades the ST-segment ("ischemic") response to exercise. Instead of using a single value for peak ST depression, our method integrates all ST-amplitude and slope changes that occur during the test, from onset of exercise to the end of recovery. By reflecting not only the depth of ST depression but the manner and time course by which ST changes develop and resolve, TES incorporates data that correlate with severity of coronary artery disease. Exercise ECGs of 70 patients who also had coronary arteriography were analyzed, as well as exercise records of 46 healthy volunteers. Sensitivity, specificity, predictive accuracy and correct classification rate were 85% (50 of 59), 98% (56 of 57), 98% (50 of 51) and 91% (50 of 55), respectively. Use of TES for qualitative interpretation increased sensitivity 10-15% compared with conventional criteria. TES distinguished the group of patients with three-vessel or left main coronary artery disease from those with two-vessel (p < 0.002) or one-vessel disease (p < 0.01). These differed from patients with no vessel disease (p < 0.05). TES also varied linearly when compared with angiographically determined coronary scores (p < 0.001, r = 0.71, see = 24.2). Thus, the use of TES greatly improves our ability to diagnose and quantify serially the extent of coronary artery disease and improves the accuracy of statistical statements relating to the probability of disease.

QUALITATIVE CLASSIFICATION of test results, i.e., as either positive or negative for induced myocardial ischemia, slowly has given way to the realization that the exercise response has quantitative features that reflect the extent and severity of cardiac disease and that yield prognostic information.1, 2 In the past decade several groups have begun using computer technology in an attempt to improve accuracy and to quantify stress electrocardiography. The computer not only improves signal quality and reduces the large interobserver variation in interpretation,3 but also allows use of different types of ST measurements in a variety of diagnostic algorithms that otherwise would be impossible.4 Several groups5-7 are measuring the degree of depression to improve diagnostic sensitivity and specificity.

Using such microcomputer-assisted signal processing and analysis we have extended the above approaches and have devised a new, quantitative treadmill exercise score that grades the electrocardiographic response to exercise. Based on the continuous recording and calculation of the J-point deviation and ST slope of two simultaneous leads whose signals have

From the Cardiovascular Research Institute and Department of Medicine, University of California, San Francisco, and Veterans Administration Medical Center, San Francisco, California.
Address for correspondence: Milton Hollenberg, M.D., 111C3 VA Medical Center, 4150 Clement Street, San Francisco, California 94121.
Received February 12, 1979; revision accepted August 7, 1979.
been processed by microcomputer, the score includes and weights many features of the exercise response that correlate with the extent and severity of disease and incorporates them into a single, highly reproducible index. These features include: 1) depth of J-point depression associated with horizontal or downsloping ST segment, 2) early onset of ST depression after exercise is begun, the occurrence of such changes at low heart rates and their persistence long after exercise is stopped, 3) decreased heart rate response to exercise; and 4) inability to perform exercise at more than low work loads. The establishment of this quantitative treadmill score allows us to view the electrocardiographic response to exercise as a spectrum from normality to the highest degree of abnormality and thereby to reflect and grade the severity of the myocardial ischemia induced.

Methods

Study Group

The population studied consisted of 70 male patients (ages 24-69 years, mean 53 years) with complaints of chest pain who underwent selective coronary arteriography within 1 month (usually within 1 week) of their treadmill exercise test. They were studied to confirm the diagnosis of coronary heart disease with angina pectoris or to evaluate them for coronary artery bypass surgery. Twenty-seven (39%) had a history of infarction and 18 (26%) showed definitive ECG evidence of previous myocardial infarction. Seventy-three patients (41%) were taking propranolol, 18 of them in doses greater than 80 mg/day. Only three patients were taking digoxin. No patients were in overt cardiac failure at the time of study and only 10 had normalized left ventricular (LV) end-diastolic volumes greater than normal (50-90 ml/m²). For analysis, patients were grouped by the number of major coronary arteries (left anterior descending, left circumflex, or right) that had a stenotic lesion greater than 50% of the luminal diameter. Of the 70 patients, 38 (54%) had three-vessel disease, 11 (16%) had two-vessel disease, 10 (14%) had one-vessel disease, and 11 (16%) had no significant disease in any major vessel. Ten patients (14%) had significant left main coronary artery disease. The mean coronary score, ejection fraction, LV end-diastolic volume and pressure for those groups of patients with 3, 2, 1, or no coronary arteries involved are shown in table 1. The relatively high mean LV end-diastolic pressure in the group with no significant stenotic lesion suggests that these patients may not be completely normal. Fifteen of the 70 patients had resting J-point depression > 0.3 mm and were included for analysis. Forty-six young, healthy volunteers in their late teens and twenties who were presumed to be free of significant coronary artery disease were exercised. Patients on drugs were included for analysis, whereas those with conduction defects or valvular heart disease were excluded from the study.

Procedures

Graded exercise testing during the fasting state was performed on a treadmill using the standard Bruce protocol. All tests were performed under direct supervision of a physician. Silver/silver-chloride electrodes were placed in the infraclavicular areas bilaterally and on or just above both iliac crests. Standard chest leads (V1-V6) were positioned and a mesh surgical body stocking was placed over the chest and upper abdomen to reduce movement of the electrodes, leads, and electrode cable. Careful attention was given to proper electrode fixation. A full, 12-lead ECG was recorded on a three-channel recorder* while the patient was at rest in the supine and then erect positions, after hyperventilating for 30 seconds while erect, and at the termination of the study. During the entire period of exercise and recovery three leads (V1, V5, and aVF) were monitored simultaneously and continuously. Exercise was continued until the patient stopped because of moderately severe chest pain, exhaustion, shortness of breath, leg pain, dizziness, or until the physician terminated the test because of frequent premature ventricular complexes (PVCs), ventricular tachycardia (three or more consecutive PVCs), hypotension, or 4-mm ST depression.

Selective coronary arteriograms were performed in multiple projections. Significant coronary arterial narrowing was defined as greater than 50% reduction in luminal diameter of one of the three major coronary arteries. The extent and severity of coronary arterial obstruction was also scored by the method of Friesinger et al., in which each of the three major

<table>
<thead>
<tr>
<th>Vessels with &gt;50% reduction of luminal diameter</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronary score</td>
<td>11.3 ± 0.3†</td>
<td>8.2 ± 0.5†</td>
<td>3.7 ± 0.6†</td>
<td>0.7 ± 0.3†</td>
</tr>
<tr>
<td>Ejection fraction</td>
<td>61 ± 3*</td>
<td>75 ± 6</td>
<td>68 ± 5</td>
<td>73 ± 2</td>
</tr>
<tr>
<td>LV end-diastolic volume (ml/m²)</td>
<td>71 ± 4</td>
<td>74 ± 7</td>
<td>73 ± 9</td>
<td>58 ± 9</td>
</tr>
<tr>
<td>LV end-diastolic pressure (mm Hg)</td>
<td>11.9 ± 0.8</td>
<td>11.1 ± 0.8</td>
<td>12.3 ± 1.7</td>
<td>13.7 ± 1.8</td>
</tr>
</tbody>
</table>


Abbreviation: LV = left ventricular.
arteries was given a score of from 0 to 5, the highest total score possible being 15. Left ventriculograms were performed in the 30° right anterior oblique position and ejection fraction, LV end-diastolic pressure and regional wall motion were measured.

Derivation of the Treadmill Exercise Score

The J-point deviation and corresponding ST-segment slope in modified V1, V5 and aVF were recorded continuously throughout the entire test from onset of exercise to end of recovery (fig. 1), at which time the ECG had returned to the resting, preexercise state. If the patient’s ECG had not returned to the baseline by 10 minutes after cessation of exercise the recovery period was terminated at this point. To derive the treadmill exercise score the individual curves that describe the changes of both J-point deviation and ST-segment slope with time for aVF and V5 over the course of the entire test were magnified 5.2 times and traced. The areas under these curves were then summed algebraically using a Hewlett-Packard digitizer (H-P 9864A) and calculator (H-P 9830A) (fig. 2). From this final summed curve the area is obtained that relates to the total integrated area of ST-segment depression for each complex throughout the entire test. This value for ST depression is then divided by the duration of exercise in minutes and by the percent of the maximal predicted heart rate achieved to yield a treadmill exercise score. Thus,

Treadmill exercise score =

\[
\text{Area of J point and ST slope curves (V}_5 + aV_F) \times \frac{\text{Duration of exercise}}{\text{Predicted maximal heart rate}}
\]

As the severity of disease increases, the J point becomes more depressed and the ST segment becomes less upsloping and more downsloping; both sets of changes appear earlier after the onset of exercise and persist longer after cessation of exercise. These factors increase the negative area and the value of the numerator and increase the treadmill exercise score. Moreover, as severity of disease increases, not only does the numerator increase, but the value of the denominator decreases, i.e., the duration of exercise and the percent of the maximal predicted heart rate decrease. These factors tend to increase further the numerical value of the score in the severely diseased patient. In patients with resting ST changes the baseline for determining the area under the curves was not the zero deviation line, but the value of deviation that existed after the period of erect hyperventilation and at the onset of exercise. Thus, the area contributed to by resting J point and ST-slope deviation was excluded from the final summed curve.

**Figure 1.** Final trend report generated by CASE system. J point and ST-slope deviation are trended in V1, V5, and aVF, from onset of exercise until termination of the recovery period. Also plotted in the right column are heart rate, frequency of premature ventricular complexes and blood pressure.
by guest on April 16, 2017 http://circ.ahajournals.org/ Downloaded from

severely diseased patients the primary indication for termination of exercise was development of chest pain (74% of patients with three-vessel disease). In patients with little or no coronary obstructive disease the major factor limiting continuation of exercise was dyspnea and exhaustion. Sixteen of 59 patients with coronary disease experienced no pain during the treadmill test. Only 20% (10 of 50) of those with true-positive tests did not experience pain, whereas 67% (six of nine) with false-negative tests by treadmill exercise score, were pain-free during the test ($p < 0.005$). The coronary score was similar in both groups of patients.

Treadmill Exercise Score

The treadmill exercise scores obtained in patients grouped by the number of coronary arteries involved — i.e., arteries with a reduction in luminal diameter of greater than 50% — are shown in figure 3. The group of young volunteers who did not undergo coronary arteriography but who were presumed to be free of disease are plotted separately. We set $-4$ as the score that divides positive from negative tests because only two of 38 patients with three-vessel disease had a false-negative test and only one of 57 persons with no significant coronary arterial involvement had a false-positive test by this criterion. For the 38 patients with three-vessel disease the mean score was $-35 \pm 5$. The treadmill exercise score could not distinguish these patients from the 10 patients with three-vessel disease who additionally had left main coronary artery disease. The mean score of patients with three-vessel disease differs significantly from that of patients with two-vessel disease ($-13 \pm 5, p < 0.002$) and one-vessel disease ($-6 \pm 3, p < 0.01$). While the treadmill exercise scores of patients with two- and one-vessel disease do not differ from each other, they do differ significantly from that of the group with no vessels involved ($+4 \pm 2, p < 0.05$). These in turn differ from that of the young, healthy volunteers ($+13 \pm 1, p < 0.01$). The mean ages for both groups with no coronary artery disease are 49 and 23 years, respectively. Their different ages and treadmill exercise scores suggest that they reflect different subpopulations.

Tables 3 and 4 show the effects of propranolol on mean treadmill exercise scores for patients with three-, two-, or one-vessel disease or no disease. In patients with three-vessel disease (table 3) treadmill scores were indistinguishable in patients taking no propranolol or less than 80 mg/day ($-38 \pm 9$ vs $-37 \pm 10$), while they tended to be lower in patients taking more than 80 mg/day ($-28 \pm 8$) ($p = NS$). Propranolol appeared not to account for the false-negative tests, since only four of nine patients with such tests were taking propranolol. Thus, propranolol in doses of 80 mg/day or less appears not to alter significantly group treadmill exercise scores.

The appearance of chest pain during the exercise test was a significant variable. Patients with coronary artery disease who developed chest pain had a mean treadmill exercise score of $-32 \pm 5$ compared with a

![Figure 2. Summation of curves obtained from $V_s, aV_y$ from record shown in figure 1.](image)

Statistical Analysis

Data are expressed as the mean ± SEM. Statistical significance of differences between groups of patients were analyzed by analysis of variance. The relationship between treadmill exercise score and the coronary artery score was evaluated by linear regression. Chi-square and unpaired and paired $t$ tests were also used.

Results

Clinical Performance during Treadmill Exercise

No complications occurred during exercise testing. Table 2 summarizes the standard performance characteristics during graded treadmill exercise testing. Patients with three-vessel disease performed less work on the treadmill than patients with one-vessel ($p < 0.05$) or no vessels involved ($p < 0.001$). All patient groups achieved similar normalized heart rates at peak exercise (expressed as percent of the maximal predicted for age) when the patients taking propranolol were excluded from analysis. In the most
TABLE 2. Treadmill Exercise Test Performance

<table>
<thead>
<tr>
<th>Number of patients</th>
<th>Vessels with 50% reduction of luminal diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Number of patients</td>
<td>38</td>
</tr>
<tr>
<td>Mean age (years) (range)</td>
<td>55 (46-69)</td>
</tr>
<tr>
<td>Exercise duration</td>
<td></td>
</tr>
<tr>
<td>All patients</td>
<td>5.1 ± 0.4†</td>
</tr>
<tr>
<td>Patients on no propranolol</td>
<td>5.3 ± 0.6‡</td>
</tr>
<tr>
<td>% maximal predicted heart rate</td>
<td></td>
</tr>
<tr>
<td>All patients</td>
<td>69 ± 2</td>
</tr>
<tr>
<td>Patients on no propranolol</td>
<td>76 ± 3</td>
</tr>
<tr>
<td>Test terminated due to:</td>
<td></td>
</tr>
<tr>
<td>Chest pain/ST depression</td>
<td>28 (74%)</td>
</tr>
<tr>
<td>Fatigue and dyspnea</td>
<td>8</td>
</tr>
<tr>
<td>Leg pain</td>
<td>2</td>
</tr>
<tr>
<td>Pain during test</td>
<td>30 (79%)</td>
</tr>
<tr>
<td>Ventricular arrhythmias</td>
<td>4 (11%)</td>
</tr>
</tbody>
</table>

*p < 0.001 when compared with any group.
†p < 0.05 when compared with patients with one vessel involved and < 0.001 when compared with patients with no vessels involved.
‡p < 0.01 when compared with patients with no vessels involved.

The treadmill exercise score was very reproducible, as demonstrated in 16 patients who had a repeat test done within 3 months of the initial test (fig. 4A). Initial and repeat exercise tests had a correlation coefficient of 0.96, and a linear regression equation whose slope varied little from the line of identity (y = 0.6 + 0.9x, p < 0.0001, SEE = 2.9 units). This contrasts with the greater variability obtained when duration of exercise during these same exercise tests is plotted and analyzed (fig. 4B) and which yields a correlation coefficient of 0.76 and a linear regression equation of y = 2.8 + 0.6x, p < 0.001, SEE = 2.1 minutes.

When the degree of coronary obstructive disease is graded not by number of vessels involved, but more quantitatively by the method of Friesinger et al., on a scale of 0–15 on the abscissa and then plotted against treadmill exercise score on the ordinate, a fairly good correlation is obtained despite moderate variability (y = 10.8 – 3.8x, p < 0.001, r = 0.71, SEE = 24.2 units). Such correlation is significant, although the score of –11 ± 3 for those who did not (p < 0.02).

![Treadmill Exercise Score](image)

**NUMBER OF CORONARY ARTERIES**

**FIGURE 3.** Treadmill exercise scores of 70 patients grouped by number of coronary arteries involved and of 46 volunteers. The dashed line at –4 separates responses classified as normal (above line) from abnormal (below line) in order to calculate test sensitivity, specificity, etc. *p < 0.002 (three- vs two-vessel disease) and < 0.01 (three- vs one-vessel disease). †p < 0.05 (two- and one-vessel vs no disease). ‡p < 0.01 (no disease vs volunteers).

TABLE 3. Treadmill Exercise Scores: Patients with Three-Vessel Disease

<table>
<thead>
<tr>
<th>Propranolol</th>
<th>Number of patients</th>
<th>Treadmill score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>–38 ± 9</td>
</tr>
<tr>
<td>≤ 80 mg/day</td>
<td>7</td>
<td>–37 ± 10</td>
</tr>
<tr>
<td>&gt; 80 mg/day</td>
<td>11</td>
<td>–28 ± 8</td>
</tr>
</tbody>
</table>

p = NS for all group comparisons.
treadmill score reflects physiologic function and presumably induced myocardial ischemia, whereas the coronary artery score is an attempt, albeit imprecise, to quantify the anatomic extent and severity of the obstructive process.

We evaluated test validity and other test characteristics by comparing the derived treadmill score with the extent of coronary artery disease as determined by coronary arteriography.

Sensitivity, or the percent of patients with disease who are correctly identified by a positive test, i.e., true positives/(true positives + false negatives) was 85% in our entire diseased population (50 of 59 patients), rose to 90% when only persons with two- or three-vessel disease were considered (44 of 49 patients), and reached 95% when persons with three-vessel disease were considered (36 of 38 patients). Sensitivity was improved significantly when the treadmill exercise score rather than conventional criteria (horizontal or downsloping ST segment with 1-mm depression 80 msec after the J point) was used to detect an abnormal test (table 5).

Specificity, or the percent of persons without disease who are correctly identified by a negative test, i.e., true negatives/(true negatives + false positives), was 91% (10 of 11 persons) and 98% (56 of 57 persons) when the young, healthy volunteers were included in the group without coronary artery disease. Again, when conventional criteria are used, specificity is lower (82%, nine of 11 persons), although the sample size is very small.

Predictive accuracy, or the percent of patients with positive tests who truly possess the disease, i.e., true positives/(true positives + false positives), was 98% (50 of 51 tests).

Correct classification rate, or the percent of all tests that are correctly classified, i.e., (true positives + true negatives)/all tests, was 86% (60 of 70 patients correctly classified as positive or negative) and 91% (106 of 116 persons if volunteers were included). The 86% correct classification rate drops to 74% (52 of 70 patients) when conventional criteria are used. These figures compare very favorably with any of the larger series reported to date (table 6).

Moreover, our method of analysis may correctly identify as negative a test that is classified as positive by conventional criteria. Thus, a patient who just achieves a 1-mm ST depression with a horizontal or downsloping ST segment for a very brief time is classified conventionally as having an abnormal response. Yet, he may have a treadmill exercise score that falls within the normal range. One such example is illustrated in figures 5 and 6. At 4 minutes of recovery 1 mm of ST depression with a slightly downsloping ST segment was achieved and was interpreted as a positive test. However, when J point

---

**Table 4. Treadmill Exercise Scores: All Groups**

<table>
<thead>
<tr>
<th>Vessels with 50% reduction of luminal diameter</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td>-35</td>
<td>-13</td>
<td>-6</td>
<td>+4</td>
</tr>
<tr>
<td>Patients taking no propranolol</td>
<td>-38</td>
<td>-17</td>
<td>-6</td>
<td>+5</td>
</tr>
<tr>
<td>Patients taking no propranolol or ≤80 mg/day</td>
<td>-38</td>
<td>-17</td>
<td>-7</td>
<td>+4</td>
</tr>
</tbody>
</table>

No significant differences exist in any particular group when treadmill exercise scores are calculated either for all patients, for only those taking no propranolol, or for those taking no propranolol or ≤80 mg/day.

*p < 0.001 when compared with all other patients, < 0.02 when compared with patients with two vessels involved, < 0.01 when compared with patients with one vessel involved, and < 0.001 when compared with patients with no vessel involved.

†p < 0.001 when compared with all other patients, < 0.05 when compared with patients with one vessel involved and < 0.005 when compared with patients with two vessels involved.

---

**Figure 4. Reproducibility of treadmill exercise test assessed by comparing the treadmill exercise score with the duration of exercise in repeat tests performed in 16 patients. A) Treadmill exercise score of repeat test plotted against that of the initial test. Line represents the line of identity. The slope of regression calculated by least squares fit is 0.9. B) Duration of exercise of repeat test plotted against that of initial test. Line represents the line of identity. The slope of regression calculated by least squares fit is 0.6.**
Conventional coronary measurements were classified as capable of being quantitatively highly reproducible. While largely empirical, his approach seemed reasonable because the sum of J-point deviation and ST slope, when values vary within the physiological range, relate fairly linearly to the area of ST depression (fig. 7). Moreover, the units of area obtained by using the equation are expressed in voltage-time, similar to Sheffield's voltage-time integral. While both variables contribute to the ST area, the slope measurement hypothetically contributes relatively less than J-point deviation. Unlike both the McHenry and Sheffield groups, who reported on measurements obtained only during peak ST deviation, our technique uses the total of all changes in ST amplitude and slope from onset of exercise to termination of recovery. The treadmill exercise score thus incorporates important information on the manner and time course by which the maximal ST amplitude and slope changes develop during exercise and resolve during recovery, i.e., how rapidly they occur and how long they persist. Thus, two patients could obtain the same maximal degree of depression of J point and/or ST slope and would have similar worst-case ST indices or ST integrals (fig. 8). However, our study indicates that patients with patterns similar to that of patient B have more extensive disease than those similar to that of patient A in figure 8. This greater severity would be indicated clearly by our method of analysis because the larger area under curve B would result in a more abnormal treadmill exercise score.

Thus, a major advantage of the treadmill exercise score is the ability to quantify and grade more precisely the entire electrocardiographic response to exercise and to identify where in the spectrum of ischemic disease from normality to the highest degree of abnormality any patient falls. Since the treadmill exercise score easily quantifies the electrocardiographic response to exercise and thereby, by inference, the severity of the myocardial ischemic response, it also lends itself particularly well to serial determinations. Thus, one could follow the progression of disease or assess the effect of medical or sur-

### Table 5. Sensitivity of Treadmill Exercise Test

<table>
<thead>
<tr>
<th>Method of analysis</th>
<th>Number of coronary arteries involved</th>
<th>Sensitivity (% of pts)</th>
<th>Specificity (% of pts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treadmill exercise score</td>
<td></td>
<td>85% (50/59)</td>
<td>90% (44/49)</td>
</tr>
<tr>
<td>Conventional (V6 and aVF)</td>
<td></td>
<td>75% (44/59)</td>
<td>82% (40/49)</td>
</tr>
<tr>
<td>Conventional (V6 only)</td>
<td></td>
<td>71% (42/59)</td>
<td>80% (39/49)</td>
</tr>
</tbody>
</table>

Depression and ST slope were considered for the entire test, a treadmill exercise score of −2 was obtained, which was within the normal range. Therefore, he was classified as negative for disease, which was confirmed by coronary arteriography.

### Discussion

The present report describes a new treadmill exercise score — a quantitative, highly reproducible index that grades the ST-segment response to exercise. Essential to the derivation of the score was the use of computer processing to eliminate excessive muscle noise, baseline drift, and other artifacts that usually result in the technically poor-quality recordings that are obtained during exercise. With the CASE system we were able to obtain technically high quality, quantifiable measurements of J-point deviation and ST slope during the period of treadmill exercise and during the recovery period. Such computer measurements also eliminated all subjectivity in deriving the score as well as the large interobserver and intraobserver variability that usually exists in conventional interpretation of the treadmill exercise test.

Using these computer measurements of ST depression similar to those used in the pioneering efforts by McHenry et al. and Sheffield and Roitman, we have extended their analyses in various ways. McHenry calculated an ST index by algebraically adding the ST amplitude and slope of the most depressed average complex (constructed from 25 cardiac cycles). Table 6 shows how the exercise test compares with coronary arteriography.

### Table 6. Exercise Test vs Coronary Arteriography

<table>
<thead>
<tr>
<th>Method of analysis</th>
<th>No. pts</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Predictive accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mason14</td>
<td>84</td>
<td>78</td>
<td>89</td>
<td>90</td>
</tr>
<tr>
<td>Kassebaum15</td>
<td>68</td>
<td>54</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>Roitman16</td>
<td>46</td>
<td>80</td>
<td>88</td>
<td>92</td>
</tr>
<tr>
<td>Ascoop17</td>
<td>96</td>
<td>59</td>
<td>94</td>
<td>89</td>
</tr>
<tr>
<td>Martin18</td>
<td>100</td>
<td>62</td>
<td>89</td>
<td>91</td>
</tr>
<tr>
<td>McHenry6</td>
<td>86</td>
<td>82</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Kelemen19</td>
<td>74</td>
<td>54</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>Bartel20</td>
<td>650</td>
<td>65</td>
<td>92</td>
<td>95</td>
</tr>
<tr>
<td>Froelicher21</td>
<td>52</td>
<td>77</td>
<td>—</td>
<td>73</td>
</tr>
<tr>
<td>Goldschlager6</td>
<td>330</td>
<td>64</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td>Hollenberg (present study)</td>
<td>70</td>
<td>85</td>
<td>91(98)</td>
<td>98</td>
</tr>
</tbody>
</table>
TREADMILL EXERCISE SCORE/Hollenberg et al.

EXHIBITED SYMPTOMS:
- CHEST PAIN
- ARM PAIN
- FATIGUE
- DIZZINESS
- ST DEPRESSION
- ARRHYTHMIA

COMMENTS:

FIGURE 5. Exercise ECG — computer-processed signals during exercise and recovery.

DURING EXERCISE

RECOVERY

FIGURE 6. Trend of J point and ST-slope deviation throughout the same test as is shown in figure 5. Note that the upgoing slope in V₅ only becomes downgoing for a very brief interval during the recovery period. Treadmill exercise score = -2.
laboratories when untrained subjects are being tested or that occur due to varying test end points chosen by different supervising physicians or technicians become much less critical to the final interpretation or score.

Use of the treadmill exercise score to identify normal and abnormal responses to exercise has yielded excellent values for test sensitivity and specificity. Sensitivity improved 10–15% over that obtained when responses were classified by conventional criteria, even when computer-processed ECG signals were used. Moreover, such high sensitivity and specificity were achieved despite inclusion of patients with QRS or ST abnormalities on the resting ECG. If the test is to be clinically useful it cannot exclude such patients, as most studies have done, since they make up a sizeable fraction of those referred. Several factors probably contribute to the enhanced sensitivity and specificity of the score in addition to the obvious power of using all computer-measured ST data from beginning of exercise to end of recovery rather than from only a single (or averaged) cardiac complex, even though the latter is computer processed. The patient with upsloping ST depression who represents a considerable dilemma to those who use conventional electrocardiographic criteria for test interpretation is no serious problem of interpretation with our technique of analysis. Some of these patients are clearly at increased risk, but good criteria to identify them do not exist. Attempts to use the slowly rising ST slope as a criterion for abnormality usually have increased sensitivity while they have decreased specificity (more false positives). Moreover, it is very difficult to quantify visually the upsloping ST segment. Computer-assisted calculation of ST slope for two simultaneous leads provides much more accurate and reproducible data than visual measurement. Also, slope is only one of the many weighted variables that determine the specific value of the score. Thus, some patients with upsloping ST segments were classified by treadmill exercise score as having mildly or moderately abnormal responses, while others were judged to be normal. Also, the variable contribution of hyperventilation to the exercise-induced ST changes was eliminated because the ECG obtained after 30 seconds of hyperventilation was used to establish the baseline from which all further J point and ST-slope deviations then were calculated. This technique probably eliminates some false-positive responders and may reduce variability in the score that is due to noncardiac factors.

However, calculation of sensitivity and specificity of treadmill exercise tests based on coronary arteriography has significant shortcomings. Coronary arteriography, which describes anatomy, cannot truly validate the treadmill exercise test, which assesses function and depends on inducing myocardial ischemia. Thus, sensitivity and specificity of the test may be even better than they appear. However, in the absence of a more appropriate standard, comparison with angiographic data is still the best corroborating method.

Despite such excellent sensitivity and specificity, however, certain limitations in the interpretation of the treadmill exercise score exist because of the nature

\[ \text{AREA} = J \cdot x + \frac{a \cdot x}{2} \]

\[ = x \left( J + \frac{a}{2} \right) \]

\[ = x \left( J + \text{slope} \cdot x \right) \]

Figure 7. Calculation of area of ST depression.

Figure 8. Trends either of J-point depression, downsloping ST segment, or summed curves depicted for two hypothetical patients. Despite similar maximal ST depression, our studies show that patient B has much more extensive coronary artery disease than patient A and probably experiences greater myocardial ischemia during exercise.
of the population studied. Relatively few persons free of significant coronary artery stenosis were studied angiographically and included in the study. Thus, the low rate of false-positive responders that we found is based on a small sample. We cannot supplement this group with the 46 young, healthy volunteers, although none of them had an abnormal score. They represent a different population than those patients in our clinic population who are referred to us for evaluation of chest pain. Moreover, no women or asymptomatic persons were included in our sample — groups that tend to decrease sensitivity and specificity when conventional criteria are used for interpretation. Our sample is further biased toward the more seriously ill patient, because those with histories of atypical chest pain who fail to achieve abnormal treadmill exercise scores usually are not investigated further with coronary arteriography in our institution. Although these types of bias are common to all studies using coronary angiographic correlation and are not unique to the present study, they do prevent application of the specific values of the treadmill exercise score to patient populations significantly dissimilar to the one studied. Moreover, the data should be verified in an independent test series.

The present study also emphasizes the significance of induced chest pain during the test. The absence of chest pain during the test predicts a significantly lower treadmill exercise score and frequently is found in patients with a negative test, i.e., both true negatives (patients without significant coronary artery disease) as well as false negatives (patients with anatomic lesions but with negative tests) (16 of 20, 80%). These findings emphasize the diagnostic importance of chest pain and give angiographic support to Cole and Elledstad's study in which they related the occurrence of angina during the test to the severity of disease. They found that men ages 41-50 years who had angina during exercise testing had a threefold greater incidence of coronary events (progression of angina, myocardial infarction, and coronary death) and a fourfold greater incidence of myocardial infarction compared with their counterparts who had ST depression alone.

Acknowledgment

We thank Harold Poehlman and Norman Honbo for their assistance in performing the treadmill exercise tests.

References

Treadmill score quantifies electrocardiographic response to exercise and improves test accuracy and reproducibility.
M Hollenberg, W R Budge, J A Wisneski and E W Gertz

Circulation. 1980;61:276-285
doi: 10.1161/01.CIR.61.2.276

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1980 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/61/2/276.citation

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/