Reproducibility of a Multi-Stage Near Maximal Treadmill Test for Exercise Tolerance in Angina Pectoris

By Paul E. Smokler, Rex N. MacAlpin, M.D., Anthony Alvaro, M.D., and Albert A. Kattus, M.D.

SUMMARY
A retrospective analysis was made of the reproducibility of a treadmill exercise tolerance test performed on patients with angina pectoris to an end point of moderately severe (3+) angina. One hundred and forty tests (70 pairs) in 63 patients were found to meet criteria for inclusion in the study. In 30 patients a comparison of the first test the patient had ever had with a subsequent test revealed a significantly poorer performance on the first test; mean differences ± one standard error of the mean between the first test and the subsequent tests were 0.90 ± 0.20 min of walking from the start of the test to the onset of angina, 2.3 ± 0.33 total min of walking, and 5.8 ± 0.86 "exercise units" attained at the end point of "3+" angina. In 40 patients a comparison of two tests, neither of which was the patients' first test, revealed no significant differences in performance; mean differences between the two tests were 0.20 ± 0.23 min of walking at the onset of angina, 0.11 ± 0.07 min of walking and 0.18 ± 0.15 exercise units attained at the end point of "3+" angina. There were no significant differences in maximum heart rates attained between any of the tests.

In using treadmill testing to quantitate exercise capacity in angina pectoris, the first test is likely to yield a poorer performance than subsequent ones. Tests subsequent to the first are highly reproducible when a specifically defined symptomatic end point of moderately severe intensity is employed.

Additional Indexing Words:
Exercise test end points

A CCURATE evaluation of disease progression and therapeutic intervention in patients with classical angina pectoris of effort requires a reliable, quantitative measure of change in functional capacity and cardiac status. Treadmill exercise tolerance testing is being used increasingly for this purpose1-5 employing a variety of end points including onset of symptoms,6,7 a target heart rate,8 and maximally tolerated effort.9 Interpretation of serial treadmill exercise tolerance tests must be done with appropriate consideration for day-to-day variation in performance unrelated to any change in the functional capacity and cardiac status of the patient. This retrospective investigation was done with the intent of determining the resolution of a multi-stage treadmill exercise tolerance test in stable classical angina pectoris patients in which the symptomatic end point was anginal pain of moderately severe (3+) intensity. The ultimate aim was to quantify the magnitude of variation in performance not due to changes in the underlying disease state and to determine the validity of a subjective but specifically defined symptomatic end point. In addition, the effect of patient familiarity with the test on serial performances was quantitatively determined.

Materials and Methods
Subjects
The exercise tolerance tests of 63 patients, 52 males and 11 females, with classical exertional angina pectoris were used in this study. All patients manifested an ischemic electrocardiographic response of horizontal or downsloping ST-segment depression of at least 0.10 mm during the test or within four minutes following completion of the exercise test.10 Coronary arteriography was done in 47 subjects and confirmed the presence of significant obstructive coronary artery disease. For the purposes of this study the patients were divided into two groups. Group I contains pairs of tests on 30 patients who had never been tested on a treadmill. Group II contains pairs of tests on 40

From the Department of Medicine, University of California School of Medicine, Los Angeles, California.
Supported by USPHS Grant HE 11823.
Address for reprints: Albert A. Kattus, M.D., Department of Medicine, Division of Cardiology, UCLA Medical Center, Los Angeles, California 90024.
Received November 20, 1972; revision accepted for publication March 26, 1973.
REPRODUCIBILITY OF TREADMILL TEST

patients who had been tested at least once prior to the first test of the pair. Seven patients met inclusion criteria for both groups.

Testing

A motorized treadmill with a 10% upgrade was used for all tests. The testing procedure has been described in detail elsewhere. Briefly, the patient begins walking at 1.0, 1.5, or 2.0 miles per hour (MPH) depending on the severity of his symptoms. The speed is increased in 0.5 MPH increments every three minutes until the patient states he is experiencing a level of angina he would grade as 3+, where 1+ is mild, 2+ is moderate, 3+ is moderately severe, and 4+ is maximally severe. In routine testing the patient is electively stopped for important arrhythmia, ataxic gait, dizziness, hypotension, 3+ dyspnea, or 3+ leg fatigue, but for this study tests stopped for reasons other than 3+ angina were excluded. The patient's description of the progression of anginal symptoms including the time of onset is recorded during the test. A bipolar, transthoracic ECG lead (RV6 to V2) is continuously monitored on an oscilloscope. ECG recordings are made during the last six seconds of every minute as well as continuously during the last six seconds of the test, the first 30 seconds of the post exercise period, and every minute in the post exercise period for ten minutes, or until the ECG returns to the pre-exercise configuration and heart rate. The patient's blood pressure is taken by the cuff method at the end of every three minute period until completion of the test. The patient is continually observed, and defibrillation and resuscitation equipment are immediately at hand. The patient is instructed not to eat for 2 hours prior to the test. The tests followed nitroglycerin by no less than 1 hour and the administration of long-acting nitrates and performance of any other exercise test by no less than 24 hours. No inquiry was made as to the proximity of the patient's most recent cigarette prior to the test. The vast majority of the subjects were non-smokers.

For this study, the paired tests were no more than six months apart. The time intervals between tests ranged from 1 to 120 days for Group I and from 1 to 154 days for Group II. Corresponding mean interval days between tests and standard error of the means were 27.2 ± 5.1 and 33.1 ± 8.1, respectively. The difference was not statistically significant. Approximately 2000 tests were reviewed from which one suitable for this analysis were culled. No tests were excluded for purely technical reasons. The vast majority of exclusions were related to major therapeutic interventions such as drug administration, surgical procedures, undertaking of exercise training programs, or to changes in disease status such as myocardial infarction, change in the resting electrocardiogram or development of prolonged chest pain without infarction. In a few cases tests were excluded because of premature stopping due to emergence of serious arrhythmia, signs of cerebral insufficiency or alteration of the format of the test. All tests used in this analysis were the first effort of the day so that the results were not influenced by the warm up phenomenon.

Method

The tests for Group I were chosen so that the first test of the pair (T1) was the patient's very first test on a treadmill and the second test (T2) was the first subsequent test satisfying all other conditions. The paired-tests included in Group II were the first pair of tests for a patient satisfying all other conditions where the first test utilized is the chronologically first test subsequent to T1 (TA) and the second test of the pair is the first acceptable test subsequent to TA (TB). One or more tests may have been performed between T1 and TA and between TA and TB, but were excluded because they failed to meet the criteria for inclusion.

The tests are summarized in terms of minutes of walking rounded off to the nearest half minute, exercise units attained, heart rate at the time of stopping, minutes of walking to the onset of angina, and heart rate at the onset of angina. (See tables 1 and 2.) The number of minutes of walking is the number of minutes the patient had walked when the 3+ angina end point was reached. The number of exercise units attained is the sum of the products of the speed of walking in MPH and the number of minutes walked at each speed. The exercise unit system of evaluation takes into account the higher work loads performed per unit of time as the speed of walking is increased. Onset heart rate is determined from an ECG recorded during the last six seconds of the minute of onset of angina and maximum heart rate is taken from the last six seconds of the test. The onset time is the time in minutes when the patient first reported a sensation of angina.

The per cent difference is the difference between the two paired-tests for a parameter divided by the value of that parameter for the first test, multiplied by 100. Levels of significance were calculated using Student's t-test for paired data.

Results

The results for Group I are summarized in table 1. A significant improvement in performance was found between the first test and the subsequent one in terms of total minutes of exercise, exercise units performed, and time to onset of angina (P ≤ 0.005). A significant difference was also found for onset heart rate (P ≤ 0.025). There were no significant differences in maximum heart rate. In a small subset of 9 subjects it was possible to compare T1 with T2 and T3 in which T2 was the second and T3 was the third acceptable test performed by the subjects. In this subset the P value for T1-T2 was ≤ 0.005 indicating a significant difference in performance, while for T2-T3 it was ≤ 0.50 indicating no significant difference.

The results for Group II are summarized in table 2. There were no significant differences in total minutes of walking, exercise units performed, maximum heart rate, onset of angina time or onset heart rate. Thus in paired tests subsequent to the
first test no significant differences of performance are shown.

Figures 1, 2 and 3 illustrate the frequency distribution of change in minutes of walking, heart rate, and angina onset time for Groups I and II. Figure 1 shows a difference in the frequency distribution of total minutes of walking between Group I and Group II with Group II approaching a normal distribution having 80% (32) of the patients within the −0.5 to +0.5 minute range, thus illustrating the high reproducibility between tests T_A and T_B. There were 17% (5 patients) within the same range in Group I showing the poor reproducibility between tests T_A and T_B. Figure 2 illustrates little difference in the frequency distribution of maximal heart rate differences with 70% of the cases falling in the −7 to +7 range for both groups. In figure 3, differences in onset time demonstrate considerable deviation from a normal distribution in both groups, indicating the wide dispersion of differences in performances if onset of angina were to be used as the end point of the test.

**Discussion**

When moderately severe (3+) angina is used as the end point for our near-maximal treadmill exercise tolerance test, it is found that highly reproducible performances are obtained on duplicate tests after the first one. The initial test usually results in a poorer performance than subsequent ones and thus might be regarded as a learning experience or an underestimation of the patient’s true functional status. Rigorous demonstration of the familiarizing effect of test number one would require that it always be compared with test number two. This was not possible because of the retrospective nature of the study; but in the small subset of 9 subjects in whom it was possible to make such comparisons the findings supported the hypothesis that further familiarization effect does not occur after the first test. Our findings on angina patients are in accord with what Erickson et al.14 found in 1946 when they measured oxygen uptake of young men walking repeatedly at 3.5 MPH on a treadmill at 10% grade. A small but consistent reduction of O_2 uptake/kg body weight occurred between the first and second runs but none thereafter. They attributed the improved efficiency on the second test to “technical training”—the

---

**Table 1**

**Statistical Data Showing Differences in Performance Between First Exercise Test and Subsequent One—Group I (N = 30)**

<table>
<thead>
<tr>
<th></th>
<th>Time to onset of angina (min)</th>
<th>Heart rate at onset of angina</th>
<th>Total minutes walked</th>
<th>Maximum heart rate</th>
<th>Exercise units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± S (T₁)</td>
<td>3.87 ± 0.47</td>
<td>*</td>
<td>5.83 ± 0.49</td>
<td>*</td>
<td>9.80 ± 1.14</td>
</tr>
<tr>
<td>Range</td>
<td>1.5 to 14</td>
<td>3.0 to 20.5</td>
<td>4.5 to 50.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± S (T₂)</td>
<td>4.77 ± 0.55</td>
<td>*</td>
<td>7.90 ± 0.66</td>
<td>*</td>
<td>14.69 ± 1.59</td>
</tr>
<tr>
<td>Range</td>
<td>2 to 14</td>
<td>3 to 20</td>
<td>4.5 to 48.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean difference</td>
<td>T₁ – T₂ = S</td>
<td>−0.90 ± 0.20</td>
<td>−4.43 ± 1.62</td>
<td>−3.04 ± 1.78</td>
<td>−4.89 ± 0.86</td>
</tr>
<tr>
<td>Range</td>
<td>−5 to 2</td>
<td>−7.0 to 3.5</td>
<td>−18.5 to 10.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>𝑃 ≤</td>
<td>0.005</td>
<td>0.025</td>
<td>0.005</td>
<td>0.10</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*Heart rates handled only as paired data. Wide range of rates makes means not meaningful.

---

**Table 2**

**Statistical Data Showing Comparison of Performances of Paired Exercise Tests Subsequent to the First One—Group II (N = 40)**

<table>
<thead>
<tr>
<th></th>
<th>Time to onset of angina (min)</th>
<th>Heart rate at onset of angina</th>
<th>Total minutes walked</th>
<th>Maximum heart rate</th>
<th>Exercise units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± S (T₂)</td>
<td>5.61 ± 0.60</td>
<td>*</td>
<td>9.27 ± 0.68</td>
<td>*</td>
<td>19.78 ± 2.04</td>
</tr>
<tr>
<td>Range</td>
<td>2 to 14</td>
<td>3.0 to 20.5</td>
<td>6.0 to 50.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± S (T₃)</td>
<td>5.81 ± 0.53</td>
<td>*</td>
<td>9.38 ± 0.67</td>
<td>*</td>
<td>19.97 ± 2.00</td>
</tr>
<tr>
<td>Range</td>
<td>2 to 14</td>
<td>3 to 20</td>
<td>4.5 to 49.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean difference</td>
<td>(T₂ – T₃) = S</td>
<td>−0.20 ± 0.23</td>
<td>2.89 ± 1.60</td>
<td>−0.11 ± 0.07</td>
<td>−0.19 ± 0.15</td>
</tr>
<tr>
<td>Range</td>
<td>−8 to 5</td>
<td>−1.5 to 1.0</td>
<td>−2.5 to 2.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>𝑃 ≤</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td></td>
<td>0.25</td>
</tr>
</tbody>
</table>

*Heart rates handled as paired data only.
patients were approximately the same for all tests suggests that the initial test may have been associated with a heart rate accelerated by anxiety.

A number of other investigators using exercise tests to study angina pectoris have relied upon the onset of anginal distress or the onset of ST-segment depression as the endpoint of their tests. Among these are Dagenais et al. who used an escalator ergometer and Blomqvist and Atkins who used a bicycle ergometer. Both investigators analyzed the reproducibility of their tests in terms of mean percent variation from their control runs. Dagenais et al. found a mean variation of 17% with no significant difference in the exercise tolerance on three control days. Blomqvist and Atkins found a 14% mean variation with intervals of one to six months between tests. Battock, Alvarez and Chidsey, using 3+ angina as their end point, found a 10.2% standard deviation between duplicate performances. Our variation between tests $T_A$ and $T_B$ expressed as average percent deviation is 5.4% for total minutes of walking time (to the nearest half minute) or 7.3% when expressed as exercise units performed.

We have chosen to evaluate the reproducibility of our test by the absolute number of minutes walked, or by the number of “exercise units” completed. This approach avoids distortion of the results by the high percentage variations produced by small changes in duration of walking when the

psychological adaptation of the subject to the unfamiliar features of the test situation. The fact that the maximum heart rates attained by our

FIGURE 1

The frequency distribution of the change in walking time to point of moderately severe (3+) angina in minutes ($T_T - T_B$ for Group I and $T_A - T_B$ for Group II). The mean change ± one standard error of the mean is $-2.07 ± 0.33$ min for Group I, and $0.11 ± 0.07$ min for Group II.

FIGURE 2

The frequency distribution of the change in heart rate at moderately severe (3+) angina ($T_T - T_B$ for Group I and $T_A - T_B$ for Group II). The mean change ± one standard error of the mean is $-3.04 ± 1.78$ for Group I, and $1.02 ± 0.81$ for Group II.

FIGURE 3

The frequency distribution of the change in time to onset of angina in minutes ($T_T - T_B$ for Group I, and $T_A - T_B$ for Group II). The mean change ± one standard error of the mean is $-0.90 ± 0.20$ min for Group I, and $-0.20 ± 0.23$ min for Group II.

Circulation, Volume XLVIII, August 1973
exercise tolerance is low and walking is done for only a few minutes. Figure 4 illustrates this point by plotting the percent deviation of our duplicate tests in Group II against the total minutes of walking on the first of the paired tests. The 20 pairs of tests in which the performances were identical are aligned along the zero line indicating no variation. Where there were variations, the highest percentages were found when the walking times were of short duration. Thus a one minute difference in walking time constituted a 20% variation when the base performance was five minutes, but a 6.7% variation when the base performance was 15 minutes.

We believe that the high degree of reproducibility of our test is accounted for by the use of a carefully considered symptomatic end point which we arrived at after several other approaches proved unsatisfactory. When our data were analyzed for reproducibility based on the endpoint of onset of anginal distress, large variations were found as shown in figure 3. Heart rates at the onset of angina showed marked variations. We therefore rejected onset of symptoms as the end point of our test.

Our arrival at the end point of 3+ intensity of distress was reached through an appreciation of strong arguments by Bruce and McDonough in favor of a maximal test in which the patient is instructed to continue exercise until he can go no further, tempered by our reluctance to subject our patients to a maximal degree of distress each time the test was carried out. Since we proposed to test each patient a number of times, we were interested in preserving patient cooperation. We were furthermore impressed with the need for a specifically defined symptomatic end point. If, indeed, we were intending to assess the severity of functional impairment imposed by the symptom of angina, then it became necessary to set the limit of exertion at some point that could be related to the limitations of function in the life experience of the person being tested. At the onset of angina or at 1+ (mild) intensity the patient is not always sure the symptom is real, and often at that time the ECG does not display the characteristic ischemic ST-segment changes. At the moderately severe or 3+ level the patient is sure of the reality of the symptoms. Because he has a firm awareness of the end point, he is able to determine it accurately and reproducibly when he performs the test another time. By the time he reaches the 3+ level the ECG changes are likely to be present if they are going to be present at all during exercise. Use of the 3+ end point avoids the variations in performance that might be attributable to the temperaments of the patients. The apprehensive and fearful patient may decide that he can go no further when he experiences the first twinge of discomfort, while the stoical patient may continue to a maximally severe or 4+ level of distress before he decides to quit.

We have not analyzed our data in terms of a 2+ end point but having decided to push the patients beyond the point of onset of symptoms, we have opted to develop a format that might be termed near maximal rather than submaximal.

Acknowledgment

The assistance of Mrs. Frances Hetz in preparing the manuscript is gratefully acknowledged.

References


Circulation, Volume XLVIII, August 1973
3. Larsen OA, Malmborg RO: Coronary Heart Disease and Physical Fitness. Baltimore, University Park Press, 1971, pp 1-278
Reproducibility of a Multi-Stage Near Maximal Treadmill Test for Exercise Tolerance in Angina Pectoris
PAUL E. SMOKLER, REX N. MACALPIN, ANTHONY ALVARO and ALBERT A. KATTUS

Circulation. 1973;48:346-351
doi: 10.1161/01.CIR.48.2.346

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/48/2/346