Clinical and Exercise Test Predictors of All-Cause Mortality*

Results From > 6,000 Consecutive Referred Male Patients

Manish Prakash, MD; Jonathan Myers, PhD; Victor F. Froelicher, MD;
Rachel Marcus, MD; Dat Do, MD; Damayanthi Kalisetti, MD; and
J. Edwin Atwood, MD

Objective: **To report the prevalence of abnormal treadmill test responses and their association with mortality in a large consecutive series of patients referred for standard exercise tests, with testing performed and reported in a standardized fashion.**

Background: **Exercise testing is widely performed, but few databases exist of large numbers of consecutive tests performed on patients referred for routine clinical purposes using standardized methods. Even fewer of the available databases have information regarding all-cause mortality as an outcome.**

Methods: **All patients referred for evaluation at two university-affiliated Veterans Affairs medical centers who underwent exercise treadmill testing for clinical indications between 1987 and 2000 were determined to be dead or alive using the Social Security death index after a mean 6.2 years (median, 7 years) of follow-up. Clinical and exercise test variables were collected prospectively according to standard definitions; testing and data management were performed in a standardized fashion using a computer-assisted protocol. All-cause mortality was utilized as the end point for follow-up. Standard survival analysis was performed, including Kaplan-Meier curves and a Cox hazard model.**

Results: **There were 6,213 male patients (mean ± SD age, 59 ± 11 years) who underwent standard exercise ECG treadmill testing over the study period with a mean follow-up duration of 6.2 ± 3.7 years. There were no complications of testing in this clinically referred population, 78% of whom were referred for chest pain, or risk factors or signs or symptoms of ischemic heart disease. Overlapping thirds had typical angina or history of myocardial infarction (MI). Five hundred seventy-nine patients had prior coronary artery bypass surgery, and 522 patients had a history of congestive heart failure (CHF). Indications for testing were in accordance with published guidelines. Twenty percent died over the follow-up period, for an average annual mortality rate of 2.6%. Cox hazard function chose the following variables in rank order as independently and significantly associated with time to death: exercise capacity (metabolic equivalents < 5, age > 65 years, history of CHF, and history of MI. A score based on these variables (summing up the four variables [if yes = 1 point]) classified patients into low-risk, medium-risk, and high-risk groups. The high-risk group (score ≥ 3) has a hazard ratio of 5.0 (95% confidence interval, 4.7 to 5.3) and a 5-year mortality rate of 31%.**

Conclusion: **This comprehensive analysis provides rates of various abnormal responses that can be expected in patients referred for exercise testing at a typical medical center. Four simple variables combined as a score powerfully stratified patients according to prognosis.**

(CHEST 2001; 120:1003–1013)

Key words: coronary artery disease; exercise testing; prognosis

Abbreviations: ACC/AHA = American College of Cardiology/American Heart Association; BMI = body mass index; CHF = congestive heart failure; EIH = exercise-induced hypotension; LBBB = left bundle branch block; LVH = left ventricular hypertrophy; MET = metabolic equivalent; MI = myocardial infarction; PVC = premature ventricular contraction; RBBB = right bundle branch block; SBP = systolic BP; VA = Veterans Affairs

Exercise testing is widely used clinically, and its many applications are the subject of a number of national[1,2] and international guidelines.[3,4] All of the guidelines relative to ischemic heart disease recommend the standard exercise test as the first choice for the evaluation of the patient who presents with chest pain and does not have resting ECG abnormalities that affect the interpretation of the exercise ECG response.[5,6] Because the literature lacks a broad picture of treadmill test utilization, it is important to present a large clinical series of consecutive patients presenting for a standard exercise test. The results can be used to provide the prevalence of unusual responses such as ST-segment elevation, exertional hypotension, bundle branch block, and ventricular tachycardia. Multicenter studies have presented treadmill results from large samples of patients. However, the patients are highly selected to fit the study protocol (Coronary Artery Surgical Study,[7] Quantitative Exercise Testing and Angiography Study Group [8]), and series from institutions (Duke,[9] Mayo,[10] and Cleveland Clinics[11]) consist of target populations to answer specific research questions. Other large databases of treadmill tests with follow-up consist of asymptomatic individuals (US Air Force,[12] Cooper Clinic,[13] Israeli Air Force[14]) rather than clinical patients. Analysis of a large, clinically referred, consecutive population also provides an opportunity to compare our indications and methods to exercise laboratories throughout the Veterans Affairs (VA) medical system. The latter is possible since we have assessed the behavior of VA exercise laboratories using questionnaires.[15] Moreover, the easy availability of Social Security death index data now readily available on the World Wide Web makes it possible to evaluate predictors of all-cause mortality. The purpose of the present study was to report the statistical distribution of treadmill test responses and their association with mortality in a large

Table 1—*Population Characteristics With Univariate Comparison Between Those Who Died and Those Who Survived**

Variables	Total Sample (n = 6,213)	Survived (n = 4,957)	Died (n = 1,256)	p Value
Demographics				
Age, yr	59 ± 11.2	57.7 ± 11.4	64.3 ± 9.1	< 0.001
Age ≥ 65 yr, %	33.8	29.3	51.4	< 0.001
Height				
Inches	69.2 ± 4.1	69.2 ± 3.8	69.2 ± 5	
cm	175.8 ± 10.4	175.8 ± 9.6	175.8 ± 12.7	0.78
Weight				
lb	191.2 ± 38.6	193.3 ± 39	183 ± 35.9	
kg	86.7 ± 17.5	87.6 ± 17.7	83 ± 16.3	< 0.001
BMI	28 ± 5.2	28.3 ± 5.3	26.7 ± 4.8	< 0.001
Inpatient, %	25.3	22.2	37.3	< 0.001
Medications, %				
Digoxin	5.4	3.9	11.5	< 0.001
Calcium antagonists	27.3	24.9	37	< 0.001
β-Blockers	18.9	18.4	21	0.03
Nitrates	23.3	19.4	38.9	< 0.001
Antihypertensives	24	21.4	34.1	< 0.001
Medical history, %				
Atrial fibrillation	3.1	2.9	4.1	0.03
Pulmonary disease	6.9	6	10.7	< 0.001
Stroke	3.6	3.2	5.5	< 0.001
Claudication	5.3	4.6	7.9	< 0.001
Typical angina	31.3	18.8	31.3	< 0.001
MI history or by Q waves	29.3	24.7	47.5	< 0.01
≤ 6 mo after MI	4.4	4.6	3.7	0.16
CHF	8.4	6.2	17.3	< 0.001
Interventions, %				
Coronary bypass surgery	9.3	7.8	15.4	< 0.001
PTCA/stents	5.2	5.3	4.5	0.28

*Data are presented as mean ± SD unless otherwise indicated. PTCA = percutaneous transluminal coronary angioplasty.

---

*From the Division of Cardiovascular Medicine, Stanford University Medical Center and the University of California Irvine, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA.

Manuscript received December 1, 2000; revision accepted February 27, 2001.

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

Table 2—*Risk Factors With Univariate Comparison Between Those Who Died and Those Who Survived**

Variables	Total Sample (n = 6,213)	Survived (n = 4,957)	Died (n = 1,256)	p Value
Diabetes, %	10.5	10.4	10.7	0.79
Hypercholesterolemia, %	29.7	31.2	23.6	< 0.001
Family history of CAD, %	22.1	23.9	14.8	< 0.001
Hypertension, %	48	47.4	50.3	0.07
Obesity (BMI > 27), %	52.8	55.5	42.3	< 0.001
Smoking, %				
Currently smoking	30.6	30.5	30.9	0.81
Ever smoked	55.2	58.4	42.7	< 0.001

*CAD = coronary artery disease.

consecutive series of patients referred for standard exercise tests, with testing performed and reported in a standardized fashion.

Materials and Methods

Population

The population consisted of consecutive patients referred to two clinical exercise laboratories (Long Beach VA, from 1987 to 1991; Palo Alto VA, 1992 to 2000) directed in consistent fashion by two of the authors (V.F.F. and J.M.). Patients who were subjects in research protocols were not considered in the analyses.

Data Collection

No imaging modality was performed in conjunction with the tests; however, expired gases were measured in approximately one fifth of patients. Both laboratories had affiliations with universities and had academic medical staffs with rotating house officers and fellows. All tests were supervised directly by these individuals or by nurse practitioners; all tests were overread by two of the investigators (V.F.F. and J.M.). A thorough clinical history, listing of medications, and risk factors were recorded prospectively at the time of exercise treadmill testing using computerized forms beginning in 1987.[16,17] The forms included standard definitions of clinical conditions and exercise responses.

Exercise Testing

Patients underwent symptom-limited treadmill testing using the US Air Force School of Aerospace Medicine protocol[18] or an individualized ramp treadmill protocol.[19] Before ramp testing, the patients answered a questionnaire to estimate their exercise capacity; this allowed most patients to reach maximal exercise within the recommended range of 8 to 12 min.[20] Patients were encouraged to only use the handrails for balance. Heart rate targets were not used as an end point or to judge the adequacy of the test. Patients did not perform a cool-down walk but were placed in a supine position as soon as possible after exercise.[21] Treatment with medications was neither changed nor stopped prior to testing. After careful skin preparation, a standard 12-lead ECG was obtained before and continuously during exercise and for at least 5 min into recovery.[22]

Visual ST-segment depression was measured at the J junction and corrected for preexercise ST-segment depression while standing; ST-segment slope was measured over the following 60 ms and classified as upsloping, horizontal, or downsloping. 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. Ventricular tachycardia was defined as a run of three or more consecutive premature ventricular contractions (PVCs), as previously described.[23] BP was measured manually, and metabolic equivalents (METs) were estimated from treadmill speed and

Table 3—*Resting ECG Abnormalities With Univariate Comparison Between Those Who Died and Those Who Survived*

Variables	Total (n = 6,213)	Survived (n = 4,957)	Died (n = 1,256)	p Value
Abnormal ECG,* %	39.9	35.7	56.1	< 0.001
Anterior Q waves, %	5.5	4.1	11	< 0.001
Inferior Q waves, %	12.2	10.6	18.4	< 0.001
Lateral Q waves, %	2.8	2.1	5.7	< 0.001
Q-wave score,† mean ± SD	0.2 ± 0.47	0.17 ± 0.42	0.35 ± 0.61	< 0.001
LBBB, %	0.9	0.7	1.9	< 0.001
RBBB, %	4.4	3.6	7.5	< 0.001
Intraventricular conduction delay, %	4.9	4.2	7.4	< 0.001
LVH, %	5.2	4.4	8.4	< 0.001
Atrial fibrillation, %	1.9	1.7	3.9	< 0.05
ST-segment depression (≥ 0.25 mm), %	9.6	7.6	17.7	< 0.001

*At least one major ECG abnormality.

†Q-wave score = anterior Q waves plus inferior Q waves plus lateral Q waves.

Table 4—*Exercise Test Responses With Univariate Comparison Between Those Who Died and Those Who Survived**

Variables	Total (n = 6,213)	Survived (n = 4,957)	Died (n = 1,256)	p Value
Angina occurred, %	11.5	11	13.9	0.01
Angina reason for stopping, %	5.8	5.5	7.2	0.05
Exercise-induced ST-segment depression ($\geq$ 1 mm), %	24.8	22.4	34.5	$<$ 0.001
Silent ischemia, %	16.1	14.9	21.2	$<$ 0.001
Rate-dependent LBBB, %	1.3	0.9	4.3	$<$ 0.001
Rate-dependent RBBB, %	4.1	3	11.6	$<$ 0.001
Rate-dependent IVCD, %	2.5	2	6.7	$<$ 0.001
ST-segment elevation without Q waves, %	0.4	0.4	0.6	0.46
ST-segment elevation with Q waves, %	1.5	1	3.2	$<$ 0.001
Exercise-induced arrhythmias (frequent PVCs or ventricular tachycardia), %	8	7.2	11.6	$<$ 0.001
Exercise-induced ventricular tachycardia ($\geq$ 3 beats), %	1.4	1.3	1.8	0.22
Resting hemodynamics, mean ± SD				
Heart rate	78 ± 24	78 ± 24	79 ± 16	0.069
SBP	134 ± 22	134 ± 21	133 ± 23	0.68
Diastolic BP	82 ± 16	83 ± 16	80 ± 13	$<$ 0.001
Maximum hemodynamics				
Heart rate, mean ± SD	136.9 ± 27.8	139.3 ± 26.4	127.5 ± 31	$<$ 0.001
Maximal heart rate $<$ 85% age predicted, %	50.9	47.9	62.7	$<$ 0.001
SBP, mean ± SD	178 ± 30	179 ± 30	171 ± 32	$<$ 0.001
Maximal SBP $<$ 130 mm Hg	4.5	3.5	8.5	$<$ 0.001
Drop SBP at maximum below pretest, %	3.9	3.3	6.4	$<$ 0.001
EIH, †%	2.8	2.4	4.5	$<$ 0.001
Diastolic BP, mean ± SD	86 ± 20	86 ± 17	85 ± 27	0.24
METs, mean ± SD	8.2 ± 4.3	8.6 ± 4.4	6.5 ± 3.5	$<$ 0.001
$<$ 5 METs, %	19	16	31.1	$<$ 0.001
Borg, mean ± SD	17.1 ± 2.8	17.2 ± 2.8	17 ± 2.9	0.12

*IVCD = intraventricular conduction defect.

†Rise and drop of $\geq$ 10 mm Hg or maximal exercise SBP $\geq$ 10 mm Hg less than pretest standing.

grade.[19] Exertional hypotension was coded as either a 10-mm Hg drop in systolic BP (SBP) after a rise, or a drop of 10 mm Hg below standing pretest. An exercise SBP code was considered in which 0 = increase $>$ 40 mm Hg, 1 = 31 to 40 mm Hg, 2 = 21 to 30 mm Hg, 3 = 11 to 20 mm Hg, 4 = 0 to 10 mm Hg, and 5 = drop below standing pretest as previously defined.[24] Δs and products were not considered in the analyses. All ECG signals were digitized and stored on compact disks after being recorded (Mortara E-scribe; Milwaukee, WI, or QUEST; Burdick/Spacelabs; Milton WI).

No test result was classified as indeterminate.[25] The exercise tests were performed, analyzed, and reported per standard protocol and utilizing a computerized database (EXTRA; Mosby Publishers; Chicago, IL).[26] The textual report was automatically downloaded into the VA centralized computer database for distribution.[27]

Follow-up

The Social Security death index was used to match all of the patients using name and social security number. The index is updated weekly, and current information was used. Death status was determined as of July 2000.

Statistical Methods

Number Crunching System Software (NCSS; Salt Lake City, UT) was used for all statistical analyses after transferring the data from a database (ACCESS; Microsoft; Redmond, WA). Total (all-cause) mortality was used as the end point for follow-up for survival analysis. Censoring was not performed since data regarding subsequent interventions were not available for all patients. Survival analysis was performed using Kaplan-Meier curves to compare variables and cut-points, and the Cox hazard function was used to demonstrate which variables were independently and significantly associated with time to death. Automatic selection of variables was performed with a Z value cutoff of 2 iterations and 20 iterations. Hazard ratios were calculated along with their 95% confidence intervals.

RESULTS

Population Demographics

This male study population had a mean (± SD) height of 69 ± 4 inches, a mean weight of 191 ± 34 lb, and a mean body mass index (BMI) of 28 ± 9 kg/m^2. Patients who died were significantly older and had a lower BMI than survivors. Overall, 74% were white, 9% were Hispanic, and 12% were black. There was no significant difference in the survivorship among the different ethnicities. Average resting heart rate was 78 ± 24 beats/min, with a corresponding mean SBP of 133 ± 22 mm Hg. No significant differences in these parameters were noted between those who survived and those who died. Other relevant variables for the entire population, for those who survived, and for the 1,256 who died over the mean 6.2-year follow-up period (median, 6 years) are presented with significance levels in Table 1. There was an average annual mortality of 2.6%. Table 2 provides the risk factor data

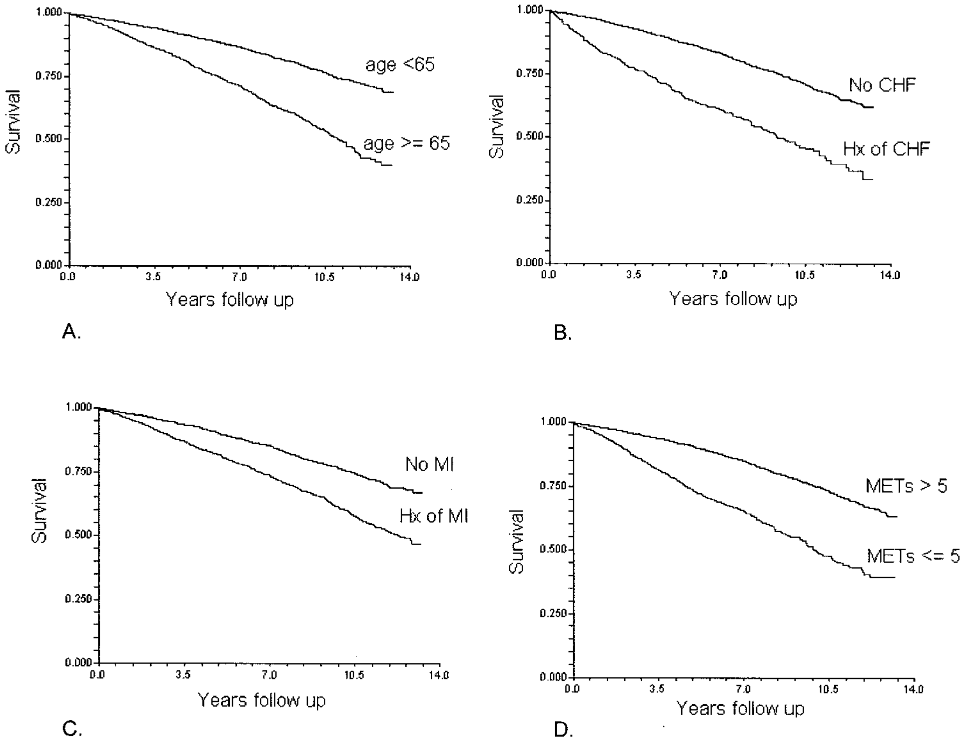

FIGURE 1. Kaplan-Meier survival curves for key clinical variables and METs. *Top left*, *A*: age < 65 years or ≥ 65 years. *Top right*, *B*: medical history of CHF. *Bottom left*, *C*: medical history of MI. *Bottom right*, *D*: exercise capacity > 5 METs or ≤ 5 METs. Hx = history.

in these groups. Risk factors for coronary disease were not clustered among those who died; in fact, risk factors trended to be clustered in those who survived. Table 3 provides data on the resting ECG. All of the major ECG abnormalities were significantly more prevalent in those who died. Medications for congestive heart failure (CHF), ischemia, and hypertension were more frequently prescribed for those who died. All of the major medical history items were significantly more prevalent in those who died, including typical angina. The prevalence of prior coronary artery bypass surgery was twice the rate in those who died compared to those who survived.

Exercise Test Responses

No major complications were encountered during testing. Results for the entire population, specifically for those who survived and the 1,256 patients who died, along with significance levels for differences are presented in Table 4. While angina occurred more frequently during the exercise test among those who died, there was only a trend for it to be more common as the reason for stopping the exercise test. Abnormal exercise-induced ST-segment depression, the most common ECG abnormality, occurred in one fourth of our population, and two thirds of the time it was silent or asymptomatic. Both symptomatic and asymptomatic ischemia (*ie*, ST-segment depression) were significantly more common in those who died. The next most prevalent ECG abnormality was frequent PVCs and/or ventricular tachycardia (three beats in a row or more), and this combined response was significantly more prevalent in those who died. ST-segment elevation was rare but more common over Q waves. ST-segment elevation was more frequent in the patients with Q waves who died, but there was no difference in those without Q waves. Exercise-induced ventricular conduction abnormalities were also rare but more common in those who died.

There were no major differences in resting heart rate or BP, but hemodynamic measurements during exercise were significantly lower in those who died. Failure to reach the age-predicted heart rate occurred in half the population and was significantly more prevalent in those who died. Exercise-induced hypotension (EIH) was relatively rare but significantly more prevalent in those who died. Both groups gave a similar effort as reflected in the mean Borg scale rating of "very hard" (Borg score of 17). The results for univariate Kaplan-Meier survival curves for selected variables are presented in Figures 1–3 and Table 5.

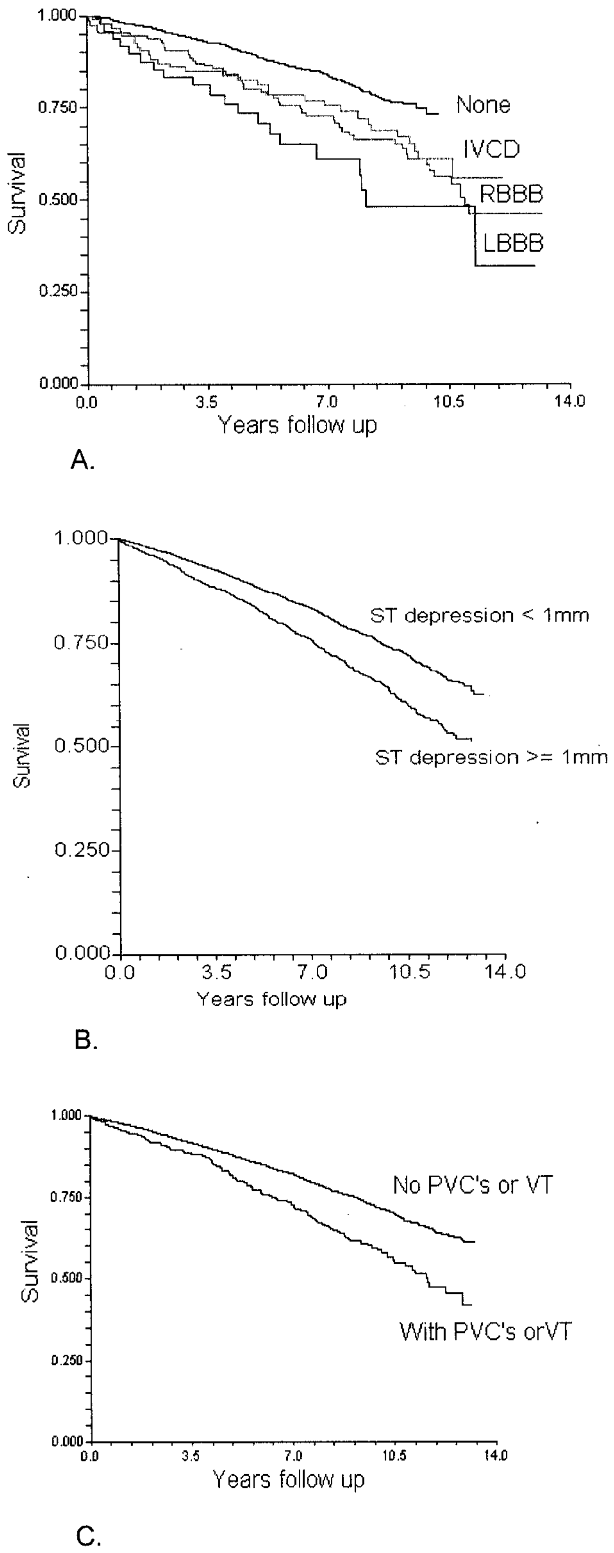

FIGURE 2. Kaplan-Meier survival curves for ECG exercise test responses. *Top, A*: exercise-induced bundle branch block. *Middle, B*: exercise-induced ST-segment depression. *Bottom, C*: exercise-induced frequent PVCs or ventricular tachycardia (VT). See Table 4 for definition of abbreviation.

Reasons for Testing

The reasons for testing are provided in Table 6. Seventy-eight percent of patients were tested for evalutation of chest pain, or evaluation of signs or symptoms possibly due to coronary disease, or for elevated risk factors. Those who survived were more likely to be tested for "softer" indications, while those who died were more likely to be tested for symptoms suggestive of an abnormality (*ie*, arrhythmia, angina).

Prognostic Score

Using stepwise selection, the proportional hazards model was allowed to build on each variable group. Table 7 lists the chosen clinical variables, and Table 8 lists the exercise variables.

Among clinical variables, age, CHF, history of MI, and pulmonary disease were significant predictors of death followed by the following ECG variables: left ventricular hypertrophy (LVH), resting ST-segment depression, resting left bundle branch block (LBBB), and resting right bundle branch block (RBBB). The most powerful exercise variables were METs and maximal heart rate. The only BP measurement chosen was an exertional hypotension coding but not a drop below standing pretest. Clinical variables alone were entered first with subsequent additions of exercise variables to arrive at the final model consisting of the following: exercise capacity (METs), age, history of MI and/or Q wave on the ECG, and history of CHF. Notably, exercise-induced ECG abnormalities were not chosen as associated with time until death. The continuous variables (age, METs) were dichotomized using cutpoints of 5 for METs and 65 years for age. When only these four variables were entered in the model, the coefficients from the Cox model were 0.67, 0.73, 0.67, and 0.50, respectively, enabling construction of the following score: METs < 5 ($1 = $ yes, $0 = $ no), plus age > 65 years ($1 = $ yes, $0 = $ no) plus history of CHF ($1 = $ yes, $0 = $ no) plus history of MI or Q wave on ECG ($1 = $ yes, $0 = $ no), as shown in Table 9. The population was then coded as to how many of these variables each subject had, and univariate survival statistics were performed using Kaplan-Meier survival curves for 0, 1, 2, and 3 or more, since so few patients had all four characteristics.

The hazard ratios, confidence intervals, and p values for score values are shown in Table 10, and the Kaplan-Meier survival curves are shown in Figure 4. The score enabled the identification of a low-risk group (42% of the cohort) with an annual mortality rate of $< 1.5\%$, two intermediate-risk groups (51% of the cohort) with an annual mortality rate from 3 to $< 5\%$, and a high-risk group (7% of the cohort) with an annual mortality rate $> 6\%$.

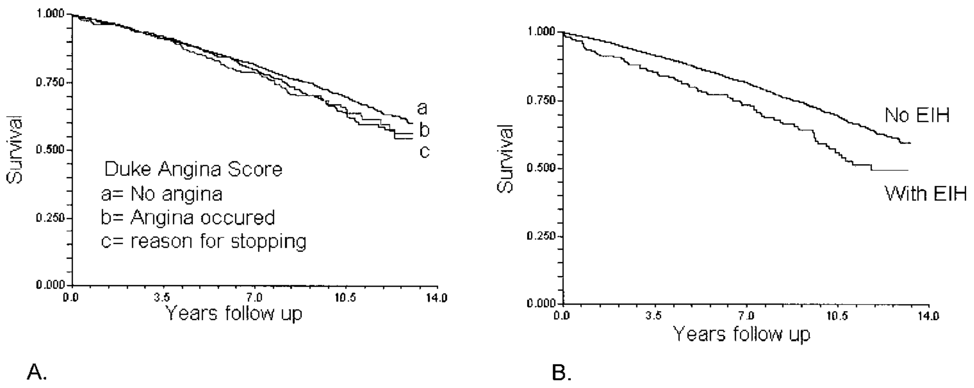

FIGURE 3. Kaplan-Meier survival curves for exercise test responses. *Left*, *A*: Duke angina score. *Right*, *B*: EIH.

DISCUSSION

The present results further define the clinical utility of the exercise test and its diagnostic and prognostic applications. The strengths of this study include an unusually long follow-up period (median, 7 years), the inclusion of all consecutive clinical referrals to two similar academically aligned VA medical centers, and the application of standardized methodology. This study is one of many to document the safety of "symptom and sign"-limited maximal exercise testing. It is important to note that in this relatively high-risk population, trained physicians or nurses were always in direct attendance of the test and one of the senior authors was always available for consultation.

Our clinical indications for the test were consistent with the American College of Cardiology/American Heart Association (ACC/AHA) guidelines[2] and were similar to exercise laboratories throughout the VA system.[15] The highest percentage of patients who underwent exercise testing in our study were tested for the diagnosis of chest pain or evaluation of possible coronary disease (78%; Table 6) fitting the ACC/AHA class 1 indication; *ie*, there is a general consensus that the tests were justified. Class 1 indications are defined as those assisting with the diagnosis of coronary artery disease, assessing exercise capacity, or estimating prognosis among patients with known heart disease, including predischarge and postdischarge testing after an MI or postrevascularization (percutaneous transluminal coronary angioplasty/stenting and coronary artery bypass surgery). However, a great deal of variation existed in terms of criteria for abnormal results and whether physician presence is required during testing at other VA medical centers.[13] We used the same ST-segment response criteria as most other VA medical centers, but only 28% of respondents to our survey used some type of treadmill score, possibly because of the lack of a score as simple as ours.

Table 5—*Univariate Comparison of Individual Variables Using Nonparametric Kaplan-Meier Statistics**

Variables	Total Patients (n = 6,213), No. (%)	Hazard Ratio	Confidence Limit (95%)	Mortality	
				5 yr	10 yr
Age ≥ 65 yr	2,097 (33.8)	2.27	2.25–2.3	20	45
CHF	522 (9.3)	1.81	1.54–2.12	30	53
MI history or Q wave on ECG	1,819 (29.3)	1.59	1.57–1.61	19	40
METs < 5	1,181 (29.3)	1.59	1.57–1.61	10	26
ST-segment depression	1,541 (24.8)	1.51	1.45–1.58	17	28
Angina occurred	717 (11.5)	1.33	1.17–1.51	13	34
Angina as a reason for stopping	602 (9.7)	1.35	1.22–1.48	20	39
Exercise-induced LBBB	48 (1.3)	2.53	1.37–4.7	30	52
Exercise-induced RBBB	151 (4.1)	1.45	1.08–1.95	21	44
Angina as a reason for stopping	375 (5.8)	1.24	0.97–1.58	NS	NS
Exercise-induced arrhythmias	503 (8)	1.70	0.70–2.04	NS	NS
EIH	245 (3.9)	1.27	0.93–1.75	NS	NS
Exercise-induced IVCD	95 (2.5)	1.22	0.78–1.89	NS	NS

*NS = not significant (confidence limit < 1); see Table 4 for definition of abbreviation.

Table 6—*Reason for Testing With Univariate Comparison Between Those Who Died and Those Who Survived*

Variables	Total (n = 6,213)	Survived (n = 4,957)	Died (n = 1,256)	p Value
Signs/symptoms/risk factors, %	34.1	38	17.8	< 0.001
Angina pectoris, %	44.1	40.7	58.2	< 0.001
Arrhythmias, %	2.9	2.5	5.1	< 0.001
Exercise capacity, %	4.3	4.7	2.5	< 0.001
Dyspnea on exertion, %	1.9	1.4	3.3	< 0.001

Our study provides additional data regarding the prevalence of a wide range of exercise test responses and their prognostic impact, and represents the findings among all consecutive clinically referred patients rather than a selected group or a research population. Table 11 provides a comparison with prior studies in regard to the frequency of the reported exercise test responses and mortality.[28] Similarities of test responses such as hemodynamics and rates of abnormal responses demonstrate that concerns about the VA not being representative of other populations are unfounded. The higher rates in prior studies for some of the variables are due to the fact that most of their populations were clinically selected to undergo cardiac catheterization (*ie*, had considerable workup bias).[28] The rates of abnormal responses, including ventricular tachycardia, ST-segment elevation, and exertional hypotension emphasize the need for proper training and prioritization of issues for certification of individuals performing exercise testing. In addition, the need for physician supervision needs to be better defined.

While many variables are univariately associated with risk for death, it is somewhat surprising that four simple variables provide the majority of the important prognostic information. Table 12 lists the number of times the major prognostic variables were chosen as significantly and independently predictive of time to death out of the times they were considered in the published prognostic studies.[24,29–36] The four most prognostic variables in the present study (age, METs, history of MI, and history of CHF) were well represented in this tabulation.

Table 7—*Variable Selection for Clinical Variables*

Variables Added	Z Value	Probability Level	χ^2
Age > 65 yr	15.86	< 0.001	231.36
CHF	11.81	< 0.001	340.28
MI and/or Q waves	9.35	< 0.001	426.53
Pulmonary disease	5.38	< 0.001	451.80
LVH	4.75	< 0.001	471.71
Resting ST-segment depression	3.71	< 0.001	486.09
Resting LBBB	3.72	< 0.001	496.75
Resting RBBB	3.60	< 0.001	508.53
Definite angina	2.34	< 0.001	514.22
Currently smoking	2.17	< 0.05	519.07

Table 8—*Variable Selection for Exercise Test Variables*

Variables Added	Z Value	Probability Level	χ^2
METs < 5	15.79	< 0.001	201
Maximal heart rate, beats/min	7.88	< 0.001	277
Exercise-induced ST-segment depression	5.75	< 0.001	308
Exercise-induced arrhythmia	5.22	< 0.001	332
Exercise-induced bundle branch block	2.83	0.005	340
Exertional hypotension	2.11	0.04	348

Other exercise variables not chosen in the final multivariate model but found to be univariately significant included maximum heart rate, exercise-induced ST-segment depression, arrhythmias during exercise, and EIH. EIH has been demonstrated in most studies to be associated with either a poor prognosis or a high risk of angiographically documented coronary disease.[37] EIH has been associated with cardiac complications during exercise testing, making measurement of BP during testing of utmost importance. In our study, EIH had a relative risk of 1.25 for all-cause mortality. Exercise-induced frequent PVCs (> 6/min) had a relative risk of 1.8, and exercise-induced ST-segment depression had a relative risk of 1.8 for all-cause mortality during the median follow-up period of 7 years. The failure of ventricular tachycardia to be independently predictive is consistent with our previous study.[23]

The relative unimportance of the ischemic variables may be due to our inability to censor on interventions for ischemia (*ie*, removal of intervened patients from observation when the intervention occurs during follow-up) and the consideration of all-cause mortality instead of cardiovascular mortality. This may also explain why the ischemic variables

Table 9—*Results of Variable Selection Considering the Best Clinical and Exercise Test Variables*

Variables Added	Z Value	Probability Level	χ^2	Regression Coefficient
METs < 5	15.90	< 0.001	204.77	0.67
Age > 65 yr	12.79	< 0.001	367.18	0.73
CHF	10.38	< 0.001	455.20	0.67
MI and/or Q waves	8.52	< 0.001	527.75	0.50

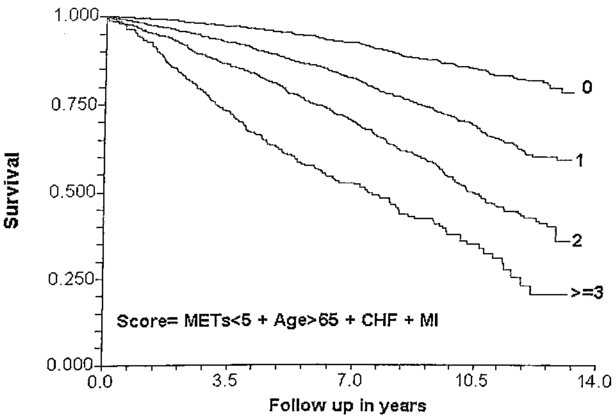

FIGURE 4. Kaplan-Meier survival curves for the score (the score is calculated by summing up the four variables [METs < 5, age > 65 years, history of CHF, and history of MI] with if yes = 1 for each).

included in the Duke score that clearly had diagnostic power[38] did not predict all-cause mortality. While all-cause mortality has advantages over cardiovascular mortality as an end point,[39] the Duke score was generated using the end points of infarction and cardiovascular death. In addition, interventions such as bypass surgery or catheter procedures were censored in the Duke study (that is, subjects were removed from the survival analysis when interventions occurred). Such censoring should increase the association of ischemic variables with outcome by removing patients whose disease has been alleviated, and thereby would not be as likely to experience the outcome. We did not censor patients based on whether they had a cardiovascular procedure during follow-up because we do not have that information. From a previous study using a similar VA patient population with an annual all-cause mortality of 3%, our group found that 75% of deaths were cardiovascular, and that 6% of patients were censored in follow-up due to bypass surgery.[40] If the proportions are similar in our current population, it would not be unreasonable to expect a bias against the predictive power of these variables. The contradictory results could also be due to the more effective methods of treatment currently available for coronary disease.

Table 10—*Values for the Score and How They Perform Prognostically*

Scores	Total Patients (n = 6,213), No. (%)	Hazard Ratio	Confidence Limit (95%)	Mortality, % 5 yr	Mortality, % 10 yr
a = 0	2,608 (42)			6	15
b = 1	2,072 (33)	2.0	2.2–2.0	12	30
c = 2	1,109 (18)	3.7	3.6–3.8	21	48
d ≥ 3	423 (7)	5.0	4.7–5.3	31	67

The use of interventions as end points falsely strengthens the association of ischemic variables with end points since the ischemic responses clinically result in the intervention being performed. While some investigators have justified their use by requiring a time period to expire after the test before using the intervention/procedure as an end point, this still influences the associations between test responses and end points. Another problem has been that variables predicting infarction can be different than those predicting death, so that variables are working against themselves when predicting infarct or death (*ie*, infarct-free survival).

What do these findings mean to the clinician? First, it should be noted that all studies have population-specific attributes that may be difficult to define. Nevertheless, if the aim is to predict infarct-free survival, the Duke treadmill score may be preferred to ours since censoring was performed and infarct-free survival was predicted. If diagnosis is the issue, either the Duke score or other treadmill diagnostic scores are indicated.[41] If diagnosis is known, prognostication using our score can help direct therapy. If diagnosis is not known and prognosis is guarded, then further diagnostic efforts may be indicated. If diagnosis is not determined and a patient is high risk by the score, then risk is likely to be improved by an exercise program and risk factor modification. If prognosis is favorable by our score, perhaps diagnosis is not as important as alleviating symptoms. Our findings strengthen the importance of exercise capacity, a reflection of the integrity of the cardiopulmonary system and a marker of a

Table 11—*Prevalence of Major Exercise Test Responses From Prior Follow-up Studies for Comparison With Our Study**

Variables	Studies, No.	Mean	Present Study
Age, mean yr	8	50	59
Exercise responses, %			
ST-segment depression	8	50	24.8
ST-segment elevation	7	3.5	1.9
Ventricular tachycardia	3	1	1.4
Frequent PVCs	8	11.6	6.6
LBBB	3	0.8	1.3
Maximum heart rate, beats/min	9	134	137
METs	8	6.5	8.2
Maximum SBP, mm Hg	8	167	178
EIH, %	5	4.6	3.9
Annual mortality, %/yr			
Cardiovascular	5	1.7	NA
All causes	5	2.2	2.6

* NA = not available/not applicable.

Table 12—*Frequency of Clinical and Exercise Test Variables Chosen as Significantly and Independently Associated With Time Until Death in Nine Previous Prognostic Studies*

Variables	Chosen, No.
Clinical	
Age	2
CHF	2
MI by history or Q waves	1
Resting ST-segment depression	1
Exercise responses	
Exercise capacity (METs)	7
Angina	5
ST-segment depression	4
Maximal heart rate	3
Maximal SBP	2
ST-segment elevation	1
PVCs	1
Maximal double product	1

physically active lifestyle, as an important predictor of survival hopefully because of, but possibly at times in spite of, modern medical treatment.

SUMMARY

This comparatively long follow-up of a large number of consecutive patients with minimal workup bias referred for exercise testing at two VA medical centers provides rates of abnormal responses that can be expected in clinical exercise laboratories adhering to the ACC/AHA indications for testing. The study has demonstrated the prognostic power of four simple pieces of information: exercise capacity, age, and history and signs of MI and CHF. These variables can be used in a simple additive score to powerfully stratify the expected risk of death after modern medical treatment. Our simple score stratified 42% of our population at low risk ($< 1.5\%$ annual mortality), 51% at intermediate risk (3 to 5% annual mortality), and 7% at high risk ($> 6\%$ annual mortality). Although ischemic ECG variables and other hemodynamic variables were univariately associated with death, they were not chosen in the hazard model as independently and significantly predictive of time until death. It may well be that modern therapies for these responses and the associated conditions are so good that they are not chosen in the hazard model. While these data are unique in terms of the size of the sample and consistency of methodology, the major limitations of only having all-cause mortality, neither data on MI as an end point nor interventional procedures for censoring, and the lack of women must be kept in mind. Nevertheless, to our knowledge, these data represent the largest series of consecutive male patients referred for a standard exercise test, all tested according to consistent and standardized methodology, and provide further evidence of the importance of exercise capacity as a determinant of prognosis.

REFERENCES

1 Fletcher GF, Balady G, Froelicher VF, et al. Exercise standards: a statement for health professionals from the American Heart Association. Circulation 1995; 91:580–615

2 Gibbons RJ, Balady GJ, Beasley JW, et al. ACC/AHA guidelines for exercise testing. J Am Coll Cardiol 1997; 30:260–311

3 Clinical exercise testing with reference to lung diseases: indications, standardization and interpretation strategies; ERS Task Force on Standardization of Clinical Exercise Testing. Eur Respir Soc Eur Respir J 1997; 10:2662–2689

4 Guidelines for cardiac exercise testing: ESC Working Group on Exercise Physiology, Physiopathology, and Electrocardiography. Eur Heart J 1993; 14:969–988

5 O'Rourke RA, Brundage BH, Froelicher V, et al. American College of Cardiology/American Heart Association expert consensus document on electron-beam computed tomography for the diagnosis and prognosis of coronary artery disease. Circulation 2000; 102:126–140

6 Braunwald E, Antman EM, Beasley JW, et al. ACC/AHA guidelines for the management of patients with unstable angina and non–ST-segment elevation myocardial infarction: executive summary and recommendations. Circulation 2000; 102:1193–1209

7 Weiner DA, McCabe CH, Ryan TJ. Prognostic assessment of patients with coronary artery disease by exercise testing. Am Heart J 1983; 105:749–755

8 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; 128:965–974

9 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

10 Kwok JM, Miller TD, Christian TF, et al. Prognostic value of a treadmill exercise score in symptomatic patients with nonspecific ST-T abnormalities on resting ECG. JAMA 1999; 282:1047–1053

11 Lauer MS, Francis GS, Okin PM, et al. Impaired chronotropic response to exercise stress testing as a predictor of mortality. JAMA 1999; 281:524–529

12 Froelicher VF, Thomas M, Pillow C, et al. An epidemiological study of asymptomatic men screened by maximal treadmill testing for latent CAD. Am J Cardiol 1974; 34:770–776

13 Gibbons, LW, Cooper K, Blair S, et al. Maximal exercise test as a predictor of risk for mortality from coronary heart disease in asymptomatic men. Am J Cardiol 2000; 86:53–58

14 Livschitz S, Sharabi Y, Yushin J, et al. Limited clinical value of exercise stress test for the screening of coronary artery disease in young, asymptomatic adult men. Am J Cardiol 2000; 86:462–464

15 Myers J, Voodi L, Umann T, et al. A survey of exercise testing: methods, utilization, interpretation, and safety in the VAHCS. J Cardiopulm Rehabil 2000; 20:251–258

16 Ustin J, Umann T, Froelicher V. Data management: a better approach. Physic Comput 1994; 12:30–33

17 Froelicher V, Shiu P. Exercise test interpretation system. Physic Comput 1996; 14:40–44

18 Wolthuis R, Froelicher VF, Fischer J, et al. New practical treadmill protocol for clinical use. Am J Cardiol 1977; 39:697–700
19 Myers J, Buchanan N, Walsh D, et al. A comparison of the ramp versus standard exercise protocols. J Am Coll Cardiol 1991; 17:1334–1342
20 Myers J, Do D, Herbert W, et al. A nomogram to predict exercise capacity from a specific activity questionnaire and clinical data. Am J Cardiol 1994; 73:591–596
21 Lachterman B, Lehmann KG, Abrahamson D, et al. "Recovery only" ST-segment depression and the predictive accuracy of the exercise test. Ann Intern Med 1990; 112:11–16
22 Gamble P, McManus H, Jensen D, et al. A comparison of the standard 12-lead electrocardiogram to exercise electrode placement. Chest 1984; 85:616–622
23 Yang JC, Wesley RC, Froelicher VF. Ventricular tachycardia during routine treadmill testing: risk and prognosis. Arch Intern Med 1991; 151:349–353
24 Morrow K, Morris CK, Froelicher VF, et al. Prediction of cardiovascular death in men undergoing noninvasive evaluation for coronary artery disease. Ann Intern Med 1993; 118:689–695
25 Reid M, Lachs M, Feinstein A. Use of methodological standards in diagnostic test research. JAMA 1995; 274:645–651
26 Shue P, Froelicher V. Extra, an expert system for exercise test reporting. J Noninvas Test 1998; II-4:21–27
27 Froelicher V, Myers J. Research as part of clinical practice: use of Windows-based relational databases. Veterans Health System J 1998; March:53–57
28 Froelicher VF, Myers J. Exercise and the heart. 4th ed. Philadelphia, PA: Saunders-Mosby, 1999
29 Mark, DB, Shaw L, Harrell FE Jr, et al. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. N Engl J Med 1991; 325:849–853
30 Gohlke H, Samek L, Betz P, et al. Exercise testing provides additional prognostic information in angiographically defined subgroups of patients with coronary artery disease. Circulation 1983; 68:979–985
31 Brunelli C, Cristofani R, L'Abbate A, for the ODI Study Group. Long-term survival in medically treated patients with ischemic heart disease and prognostic importance of clinical and electrocardiographic data.: the Italian CNR Multicenter Prospective Study OD1. Eur Heart J 1989; 10:292–303
32 Weiner DA, Ryan T, McCabe CH, et al. Prognostic importance of a clinical profile and exercise test in medically treated patients with coronary artery disease. J Am Coll Cardiol 1984; 3:772–779
33 Peduzzi P, Hultgren H, Thomsen J, et al. Prognostic value of baseline exercise tests. Prog Cardiovasc Dis 1986; 28:285–292
34 Lerman J, Svetlize H, Capris T, et al. Follow-up of patients after exercise test and catheterization. Medicina (Buenos Aires) 1986; 46:201–211
35 Wyns W, Musschaert-Beauthier E, Van Domburg R, et al. Prognostic value of symptom limited exercise testing in men with a high prevalence of coronary artery disease. Eur Heart J 1985; 6:939–945
36 Detry JM, Luwaert J, Melin J, et al. Non-invasive data provide independent prognostic information in patients with chest pain without previous myocardial infarction: findings in male patients who have had cardiac catheterization. Eur Heart J 1988; 9:418–426
37 Dubach P, Froelicher VF, Klein J, et al. Exercise-induced hypotension in a male population: criteria, causes, and prognosis. Circulation 1988; 78:1380–1387
38 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
39 Lauer MS, Blackstone EH, Young JB, et al. Cause of death in clinical research: time for a reassessment? J Am Coll Cardiol 1999; 34:618–620
40 Froelicher VF, Morrow K, Brown M, et al. Prediction of atherosclerotic cardiovascular death in men using a prognostic score. Am J Cardiol 1994; 73:133–138
41 Do D, West JA, Morise A, et al. A consensus approach to diagnosing coronary artery disease based on clinical and exercise test data. Chest 1997; 111:1742–1749