Effect of Age and End Point on the Prognostic Value of the Exercise Test*

Takuya Yamazaki, MD; Jonathan Myers, PhD; and Victor F. Froelicher, MD

Background: **The clinical and exercise test variables chosen for predicting prognosis vary in the available studies. This could be due to the effect of age of the patients tested and the choice of outcomes used as end points in these follow-up studies.**

Objective: **To evaluate the effect of age and end points on exercise test variables chosen as significantly and independently associated with time to death.**

Methods: **Analyses were performed on the first treadmill test performed on consecutive male veterans at the Palo Alto and Long Beach Veterans Affairs Medical Centers since 1987. After removal of patients with congestive heart failure, coronary interventions, left bundle-branch block, atrial fibrillation, myocardial infarction and/or Q wave, and digoxin use, 3,745 male subjects remained. The outcomes were cardiovascular and all-cause mortality. The study population was divided into subsets according to age; exercise test and clinical variables were analyzed within the age subsets using the Cox hazard model.**

Results: **The mean age at the time of testing was 57 ± 12 years (± SD) and they were followed up for a mean of 6.6 years. There were 544 all-cause deaths, with 206 of the deaths being due to cardiovascular causes (38%). When the study group was classified into subsets based on age, exercise capacity (in metabolic equivalents [METs]) was chosen by the Cox hazard model most consistently in the age groups using either end point. Even when age was added to the Duke treadmill score, prediction of death did not improve in those > 70 years of age because of the nonlinear relationship between age, the exercise test variables, and time to death. The most important age cut points for clinically important differences in exercise test predictors appeared to be 70 years and 75 years of age. In the patients 70 to 75 years of age, peak METs was the only variable predictive of all-cause mortality, and exercise-induced ST-segment depression was the only predictor of cardiovascular death; in the patients > 75 years of age, none of the exercise test responses were predictive of either death outcome.**

Conclusion: **Both age and the outcome selected as an end point affect the exercise test responses chosen for scores to predict prognosis. Differences in age of the subjects tested and/or the outcome selected as the end point can explain the differences in the studies using exercise testing to predict prognosis.**

(CHEST 2004; 125:1920–1928)

Key words: age; cause of death; exercise testing; prognosis; treadmill

Abbreviations: AUC = area under the curve; CAD = coronary artery disease; CHF = congestive heart failure; DTS = Duke treadmill score; MET = metabolic equivalent; MI = myocardial infarction; SBP = systolic BP; VA = Veterans Affairs

The standard exercise test is widely used to evaluate prognosis in patients with known or suspected coronary artery disease (CAD). The clinical and exercise test variables chosen for predicting prognosis vary in the available studies, even in those using recommended statistical techniques.[1] Surprisingly, though age is a powerful predictor of death, it has not been included in the major treadmill prognostic scores. This could be due to the age range of the patients tested, the choice of end points, or both. These issues are important since hemodynamic re-

*From the Division of Cardiovascular Medicine, Stanford University Medical Center, Veterans Affairs Health Care Systems, Long Beach and Palo Alto, CA.

Manuscript received November 14, 2003; accepted November 18, 2003.

Correspondence to: Victor Froelicher, MD, VA Palo Alto Health Care System, 3801 Miranda Ave, 111C, Palo Alto, CA 94304; e-mail: vicmd@aol.com

sponses can be related to diseases leading to noncardiovascular deaths, while the ischemic responses should be more specific for cardiovascular deaths. The purpose of this study was to determine which exercise test variables are most predictive for all-cause death or cardiovascular mortality within different age groups.

Materials and Methods

Study Population and Data Collection

Consecutive male patients (n = 6,213) referred for exercise tests for clinical indications at two university-affiliated Veterans Affairs (VA) Medical Centers (Long Beach VA from 1987 to 1991 and Palo Alto VA from 1992 to 2000) were considered for analysis. These laboratories were directed in consistent fashion by two of the authors (V.F.F. and J.M.). Both laboratories were affiliated with universities and had academic medical staffs with rotating house officers and fellows. Tests were directly supervised by these physicians or by nurse practitioners; all tests were overread by two of the investigators (V.F.F. and J.M.). No imaging modalities were performed in conjunction with exercise testing. Patients already enrolled in research protocols were not considered in the analyses. A thorough clinical history, a list of medications, and cardiac risk factors were recorded prospectively at the time of testing using computerized forms.[2,3] The forms included standard definitions of clinical conditions and exercise responses.

Patients with a history and/or findings of congestive heart failure (CHF), atrial fibrillation, left bundle-branch block, digitalis use, history of myocardial infarction (MI) and/or Q waves, and history of surgical or catheter intervention were excluded. Female patients were also excluded because of their limited numbers. The final study population consisted of 3,745 male veterans. The age divisions were performed by commonly used cut-points and then by decade: < 45, 45 to 55, 55 to 65, 65 to 70, 70 to 75, and > 75 years. The study protocol was approved by the hospital review committee at both VA medical centers. Written informed consent was obtained from all patients.

Exercise Testing

Patients underwent symptom-limited treadmill testing using a graded protocol[4] or an individualized ramp treadmill protocol.[5,6] Heart rate targets were not used as an end point or to judge the adequacy of the test. Patients were placed in supine position immediately after exercise.[7] No medications were changed or stopped prior to testing, and no test was classified as indeterminate.[8] Trained physicians or nurses were always in direct attendance of the test, and the senior authors were always available for consultation and overread all studies.

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 down sloping 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 as previously described.[9] BP was measured manually. Estimated metabolic equivalents (METs) were calculated from treadmill speed and grade.

Follow-up

The California Death registry was used to match all of the patients using name and social security number. The index is updated yearly, and current information was used. Data on subsequent interventions or nonfatal cardiovascular events were not available. Death status was determined as of July 2000.

Statistical Methods

Statistical analyses were performed (Number Cruncher Statistical Systems Software; NCSS; Salt Lake City, UT) after transferring the data from an database (ACCESS; Microsoft; Redmond, WA). Patients were not removed from observation at the time of interventions or infarction since this information was not available. Clinical and exercise variables of those who survived and those who died (all-cause death and cardiovascular death) were compared using χ^2 tests for categorical variables and unpaired *t* tests for continuous variables. The study sample was divided into age groups, and the Cox hazard function was used to demonstrate which variables were independently and significantly associated with time to death using a Z value cutoff of 2, with a maximum iteration of 20.

Score Derivation

The Duke treadmill score (DTS) was calculated as: exercise time − 5 × (amount ST-segment depression) − 4 × (treadmill angina), with 0 = none, 1 = nonlimiting, and 2 = exercise-limiting angina. The METs achieved by each patient exercising on the individualized ramp protocol were converted to exercise time per the Bruce protocol and inputted into the DTS equation. The DTS and age were entered in the Cox hazard model. In addition, various mathematical transformations of age were entered into the model. Using receiver operating characteristic analysis with cardiovascular mortality as the outcome, AUCs were calculated to compare the DTS and DTS − age. A z-score was calculated to determine if the differences between the areas under the curve (AUCs) were statistically significant ($Z > 1.96 = p < 0.001$).

Results

Population Demographics, Hemodynamic and ECG Responses

Mean follow-up was 6.4 years, during which time 544 all-cause deaths were observed, 206 of which had a cardiovascular cause (38%). Annual mortality was 0.7% for cardiovascular death and 1.8% for all-cause deaths. There were significant differences for nitrate and antihypertensive drug use between the survived and death groups but not for beta-blocker use (Table 1). Those who died were significantly more likely to have pulmonary disease, claudication, typical angina, and hypertension. No significant differences were found regarding diabetes. Resting ST-segment depression was more common and resting systolic BP (SBP) was significantly higher in those who died. Angina occurred during testing in 12.1% of the population, and it was the reason for stopping in 3.7%; both were more prevalent in those who died. Abnormal exercise-induced

Table 1—*Demographic and Clinical Characteristics of Surviving, Cardiovascular Death, and All-Cause Death**

Characteristics	Total Subjects (n = 3,745)	Survived (n = 3,200)	Cardiovascular Death (n = 206)	p Value	All-Cause Death (n = 544)	p Value
Age, yr	57.3 ± 11.6	56.3 ± 11.5	63.7 ± 9.5	< 0.0001	63.2 ± 9.7	< 0.0001
Medication use, %						
Nitrates	497 (13.3)	372 (11.6)	52 (25.2)	< 0.0001	125 (22.9)	< 0.0001
β-blocker	507 (13.5)	426 (13.3)	37 (18.0)	0.06	81 (14.9)	0.34
Antihypertensives	814 (21.7)	639 (20.0)	85 (41.3)	< 0.0001	175 (32.1)	< 0.0001
Medical history and cardiac risk factors						
Claudication	151 (4.0)	119 (3.7)	19 (9.2)	< 0.0001	32 (5.9)	0.02
Stroke	101 (2.7)	77 (2.4)	9 (4.4)	0.08	24 (4.4)	0.01
Pulmonary disease	226 (6.0)	168 (5.3)	20 (9.7)	0.006	58 (10.6)	< 0.0001
Typical angina	496 (13.2)	386 (12.1)	45 (21.8)	< 0.0001	110 (20.2)	< 0.0001
Hypertension	1,727 (46.1)	1,451 (45.3)	128 (62.1)	< 0.0001	276 (50.6)	0.02
Diabetes	358 (9.6)	297 (9.3)	26 (12.6)	0.11	61 (11.2)	0.18
Resting values						
RBBB	123 (3.3)	91 (2.8)	9 (4.4)	0.21	32 (5.9)	< 0.0001
IVCD	127 (3.4)	103 (3.2)	13 (6.3)	0.02	24 (4.4)	0.16
LVH	172 (4.6)	128 (4.0)	25 (12.1)	< 0.0001	44 (8.1)	< 0.0001
ST-segment depression, mm	215 (5.7)	152 (4.8)	31 (15.0)	< 0.0001	63 (11.6)	< 0.0001
< 1	192 (89.3)	138 (90.8)	26 (83.9)	0.24	54 (85.7)	< 0.0001
≥1	23 (10.7)	14 (9.2)	5 (16.1)	0.002	9 (14.3)	0.63
Heart rate, beats/min	77.9 ± 14.9	77.4 ± 14.6	79.0 ± 14.1	0.13	80.9 ± 16.4	< 0.0001
Systolic BP, mm Hg	134.6 ± 20.4	134.2 ± 20.0	139.2 ± 21.9	0.002	136.8 ± 22.8	0.006
Exercise values						
Exercise-induced ST-segment depression	695 (18.6)	544 (17.0)	72 (35.0)	< 0.0001	151 (27.7)	< 0.0001
Silent ischemia	461 (12.3)	372 (11.6)	42 (20.4)	< 0.0001	89 (16.3)	0.002
Exercise angina	454 (12.1)	369 (11.5)	38 (18.4)	0.003	85 (15.6)	0.009
Maximal heart rate, beats/min	142.2 ± 23.6	144.0 ± 23.3	130.8 ± 22.3	< 0.0001	131.7 ± 22.5	< 0.0001
Maximal SBP, mm Hg	183.5 ± 27.6	184.1 ± 28.2	183.9 ± 27.3	0.93	180.7 ± 29.2	0.013
METs	8.9 ± 3.8	9.2 ± 3.7	7.2 ± 3.0	< 0.0001	7.1 ± 3.3	< 0.0001
Duke angina index	0.18 ± 0.47	0.17 ± 0.45	0.28 ± 0.58	0.0005	0.23 ± 0.54	0.002
Index = 0	3,227 (86.2)	2,779 (74.2)	162 (78.6)	0.0009	448 (82.2)	0.004
Index = 1	379 (10.1)	312 (8.3)	30 (14.6)	0.03	67 (12.3)	0.08
Index = 2	139 (3.7)	109 (2.9)	14 (6.8)	0.01	30 (5.5)	0.02

*Data are presented as mean ± SD or No. (%). RBBB = right bundle-branch block; IVCD = intraventricular conduction defect; LVH = left ventricular hypertrophy.

ST-segment depression and silent ischemia (*ie*, ST-segment depression without chest pain) were more prevalent among those who died compared to survivors (11.6% and 20.4%, respectively). The mean Duke angina index was significantly higher in both death groups compared to survivors (p = 0.0005).

Multivariate Analyses Using Cardiovascular Mortality and All-Cause Mortality as End Points

Table 2 shows the performance of the exercise test responses and age from the Cox proportional hazards model for predicting cardiovascular and all-cause

Table 2—*Exercise Test Variables Chosen in the Cox Proportional Hazards Model as Significantly and Independently Associated With Time Until Death Due to Cardiac and All Causes in the Entire Population**

Variables	Cardiovascular Death, Hazard Ratio (95% CI)	p Value	All-Cause Death, Hazard Ratio (95% CI)	p Value
METs	0.91 (0.86–0.96)	< 0.0001	0.89 (0.86–0.92)	< 0.0001
Age	1.03 (1.02–1.05)	< 0.0001	1.03 (1.02–1.04)	< 0.0001
Exercise-induced ST-segment depression	1.26 (1.10–1.43)	< 0.001	NS	NS
Rest ST-segment depression	1.6 (1.05–2.36)	0.03	1.29 (0.98–1.70)	0.04
Duke angina index	NS	NS	NS	NS
Maximal SBP	NS	NS	NS	NS
Maximal heart rate	NS	NS	NS	NS

*CI = confidence interval; NS = not significant.

death in the entire study population. METs and age were chosen as independently and significantly associated with time to both cardiovascular and all-cause death. Resting ST-segment depression was weakly selected for both end points, but exercise-induced ST-segment depression was chosen only by the model using cardiovascular death as the end point. Each increase of 1 MET was associated with an 11% decrease in risk for all-cause death and a 9% decrease for cardiovascular death. The Duke angina index, maximal SBP, and maximal heart rate were also considered in the model, but none of them were associated with time to either cardiovascular or all-cause death.

Treadmill Responses and Multivariate Analyses in Different Age Divisions

The descriptive statistics for the exercise test responses and survival within each age division are shown in Table 3. As expected, survival declined with age and the relative percentage of deaths due to cardiovascular causes was constant at approximately 40%. METs and maximal heart rate declined with age, while maximal SBP was relatively constant. The prevalence of exercise-induced ST-segment depression and Duke treadmill angina index were the lowest in the youngest subjects and increased with age. There were minimal differences in the mean Borg scale or in reasons for stopping across the age groups.

The results of Cox proportional hazard analysis within each age-division subsets are shown for predicting cardiovascular death and for all-cause death are shown in Table 4. The three components of the DTS (*ie*, METs, exercise-induced ST-segment depression, and Duke treadmill angina index), maximal SBP, and heart rate were considered in the model. Resting ST-segment depression was also included since it can impact the results of exercise testing, but neither it nor maximal heart rate were chosen. Using cardiovascular death as the outcome, METs were selected first in each age-division between 45 years and 70 years of age (Table 4). The Duke angina index was not selected as a significant predictor in any age group. Exercise-induced ST-segment depression was the strongest predictor in those 45 to 55 years of age and in those 70 to 75 years of age. None of the variables were chosen as predictive in the youngest group (< 45 years of age) with the lowest of cardiovascular mortality (1.1%), or the oldest group (> 75 years of age) with the highest cardiovascular mortality (10.1%).

For all-cause death in Table 4, METs were chosen as the most important predictor except in those ≥ 75 years of age where no variable was predictive. Exer-

Table 3—*Distribution of Treadmill Variables Within Age Divisions**

Variables	Age, yr							
	< 45	45–55	< 55	55–65	> 65	65–75	70–75	≥ 75
Patients, No.	624	1,014	1,638	1,138	966	778	319	188
Cardiovascular death	7 (1.1)	35 (3.5)	42 (2.6)	69 (6.1)	94 (9.7)	75 (9.6)	37 (11.6)	19 (10.1)
All-cause death	33 (5.3)	77 (7.6)	110 (6.7)	198 (17.4)	236 (24.4)	190 (24.4)	83 (26.0)	46 (24.5)
% Deaths CV, %†	21	45	38	35	40	40	45	41
Resting ST-segment depression	9 (1.4)	36 (3.6)	45 (2.7)	82 (7.2)	88 (9.1)	64 (8.2)	30 (9.4)	24 (12.8)
METs	11.9 ± 3.9	9.7 ± 3.3	10.6 ± 3.7	8.3 ± 3.4	6.8 ± 2.9	7.0 ± 3.0	6.5 ± 2.8	5.6 ± 2.5
Exercise-induced ST-segment depression, mm								
> 1.0	40 (6.4)	130 (12.8)	170 (10.4)	261 (22.9)	264 (27.3)	208 (26.7)	88 (27.6)	56 (29.8)
1.0 to 1.9	23 (3.6)	72 (7.1)	95 (5.8)	123 (10.8)	130 (13.5)	106 (13.6)	47 (14.7)	24 (12.8)
> 2.0	17 (2.7)	58 (5.7)	75 (4.6)	138 (12.1)	134 (13.9)	102 (13.1)	41 (12.9)	32 (17.0)
Duke angina index	0.11 ± 0.37	0.15 ± 0.43	0.13 ± 0.41	0.19 ± 0.48	0.23 ± 0.54	0.22 ± 0.54	0.21 ± 0.50	0.23 ± 0.54
Index-0	569 (91.2)	889 (87.7)	1,458 (89.0)	962 (84.5)	804 (83.2)	648 (83.3)	267 (83.7)	156 (83.0)
Index-1	42 (6.7)	97 (9.6)	139 (8.5)	134 (11.8)	106 (11.0)	85 (10.9)	38 (11.9)	21 (11.2)
Index-2	13 (2.1)	28 (2.8)	41 (2.5)	42 (3.7)	56 (5.8)	45 (5.8)	14 (4.4)	11 (5.9)
Maximal systolic BP, mm Hg	181.2 ± 27.4	181.9 ± 27.4	181.6 ± 27.4	186.0 ± 28.0	183.7 ± 27.3	184.7 ± 27.6	182.8 ± 26.0	179.4 ± 25.9
Maximal heart rate, beats/min	158.4 ± 22.1	147.5 ± 21.1	151.7 ± 22.1	138.0 ± 21.5	131.1 ± 22.0	133.2 ± 22.4	128.5 ± 21.7	122.4 ± 17.8
Borg scale	17.4 ± 2.2	17.4 ± 2.1	17.4 ± 2.1	17.5 ± 2.2	17.2 ± 2.4	17.3 ± 2.4	17.2 ± 2.3	16.9 ± 2.2
Reason for stopping‡								
Fatigue	276 (44.2)	589 (58.1)	865 (52.8)	409 (35.9)	403 (41.7)	292 (37.5)	131 (41.1)	111 (59.0)
Claudication	6 (1.0)	34 (3.4)	40 (2.4)	70 (6.2)	53 (5.5)	46 (5.9)	18 (5.6)	7 (3.7)
Leg pain	59 (9.5)	112 (11.0)	171 (10.4)	151 (13.3)	124 (12.8)	109 (14.0)	44 (13.8)	15 (8.0)
Submaximal test	41 (6.6)	85 (8.4)	126 (7.7)	184 (16.2)	173 (17.9)	150 (19.3)	64 (20.1)	23 (12.2)

*Data are presented as No. (%) or mean ± SD.
†% Deaths CV = percentage of cardiovascular events.
‡Patient's reason for stopping during exercise testing.

Table 4—*Exercise Test Variables Chosen in the Cox Proportional Hazards Model**

	Age, yr											
	< 45 (n = 624)		45 to 55 (n = 1,014)		55 to 65 (n = 1,138)		65 to 75 (n = 778)		70 to 75 (n = 319)		≥ 75 (n = 188)	
Cardiovascular death, no. (%)	7 (1.1)		35 (3.5)		69 (6.1)		75 (9.6)		37 (11.6)		19 (10.1)	
METs	(−)	NS	0.83 (0.75–0.93)‡	$p = 0.003$	0.88 (0.82–0.95)†	$p = 0.001$	0.88 (0.82–0.97)†	$p = 0.002$		NS	(−)	NS
Exercise-induced ST-segment depression	(−)	NS	1.84 (1.38–2.46)†	$p < 0.0001$	(−)	NS	(−)	NS	1.52 (1.15–2.01)	$p = 0.002$	(−)	NS
Maximal SBP	(−)	NS	1.01 (1.00–1.03)§	$p = 0.02$	(−)	NS	(−)	NS		NS	(−)	NS
All-cause death, No. (%)	33 (5.3)		77 (7.6)		198 (17.4)		190 (24.4)		83 (26.0)		46 (24.5)	
METs	0.86 (0.78–0.94)†	$p < 0.001$	0.89 (0.83–0.95)†	$p < 0.0005$	0.88 (0.84–0.92)†	$p < 0.0001$	0.85 (0.80–0.89)†	$p < 0.0001$	0.88 (0.80–0.95)	$p = 0.001$	(−)	NS
Exercise-induced ST-segment depression	(−)	NS	1.34 (1.06–1.68)‡	$p = 0.01$	(−)	NS	(−)	NS	1.26 (1.03–1.55)	$p = 0.03$	(−)	NS
Maximal SBP	(−)	NS	(−)	NS	0.99 (0.99–1.00)‡	$p = 0.02$	(−)	NS		NS	(−)	NS

*Data are presented as hazard ratio (95% confidence interval) unless otherwise indicated. (−)-data are missing or unavailable. See Table 2 for expansion of abbreviation. Variables entered into the model include rest ST-segment depression, METs, exercise-induced ST-segment depression, Duke treadmill angina index, maximal SBP, and maximal heart rate.

†Ranked as first in each age group.

‡Ranked as second.

§Ranked as third.

cise-induced ST-segment depression was only predictive in the 45- to 55-year-old group, and only weakly ($p < 0.01$).

Thus, only three variables, regardless of the end point, appeared in the age-specific models: METs, exercise-induced ST-segment depression, and maximal SBP. METs were most consistently chosen first except in those ≥ 75 years old where no variable was predictive. Exercise-induced ST-segment depression was chosen first only when cardiovascular death was the end point. Maximal SBP was only chosen weakly and in only two age subgroups.

Comparison of DTS and Age

Since age was not considered in the DTS, the results of an age-adjusted Cox hazard model are shown in Table 5. Both age and the DTS were chosen as independently and significantly associated with time to cardiovascular death. The hazard ratios for the DTS and age were 0.95 and 1.05, respectively ($p < 0.0001$). The DTS and age were found to have similar coefficients providing the weight of importance of the variables but opposite sign. These results were used to create an age-adjusted score, DTS − age. The original DTS and DTS − age were compared using receiver operating characteristic analysis in Figure 1. The AUC for DTS − age and original DTS were 0.70 and 0.66, respectively. DTS − age had significantly greater discriminatory power (Z-score > 1.96, $p < 0.05$) than the DTS. However, although DTS − age provided a significantly higher AUC compared to the DTS alone in the whole population and the younger patients, but there was no difference in those > 70 years of age (Fig 2).

Table 6 presents the results of the Cox hazard models with all the variables entered for those > 70 years and < 70 years of age. The salient finding was that the only treadmill variable chosen in those > 70 years old was METs when all-cause mortality was the end point, and only exercise-induced ST-segment depression when cardiovascular death was the end point. For those < 70 years old, METs and exercise-induced ST-segment depression predicted cardiovascular death, and METs and maximal heart rate predicted all-cause mortality.

Discussion

Clinical guidelines regarding ischemic heart disease have recommended the standard exercise test as the first choice for the evaluation of the elderly patient without confounding resting ECG abnormalities and that the DTS be used for prognostication.[10] The DTS was derived using infarct-free survival

Table 5—*Performance of the DTS and Age in the Cox Proportional Hazards Model for Predicting Cardiovascular Death*

Variables	Regression Coefficient	Hazard Ratio (95% CI)	Z Value	χ^2	p Value
DTS	− 0.047	0.95 (0.94–0.96)	12.9	142.2	< 0.0001
Age	0.046	1.05 (1.04–1.06)	8.6	230.1	< 0.0001

from a cohort of patients admitted for cardiac catheterization with a mean age of 49 years,[11] and later validated in outpatients (to lessen workup bias) using all-cause mortality as an end point.[12] It has been recommended by the American College of Cardiology/American Heart Association guidelines for exercise testing, and widely applied for risk stratification of patients with suspected CAD.[13]

The DTS can be considered the culmination of years of work and many studies that explored this subject often with conflicting results.[1] Different exercise responses, among the responses we considered in this study, were chosen by the various studies; surprisingly, exercise-induced ST-segment depression was not always a significant predictor. This was partially due to the fact that hemodynamic responses can be related to diseases leading to noncardiovascular deaths, while the ischemic responses are more specific for cardiovascular deaths. While prior studies[1] usually included exercise test-induced angina, the DTS added a score that graded angina according to occurrence (= 1) and being the reason for stopping (= 2). These older studies were further complicated by workup bias and by including patients with CHF who are known to have other predictive variables, particularly those relating to ventricular function.[14] It is even questionable now whether such follow-up studies can be performed since the interventions so greatly alter the natural history of CAD.

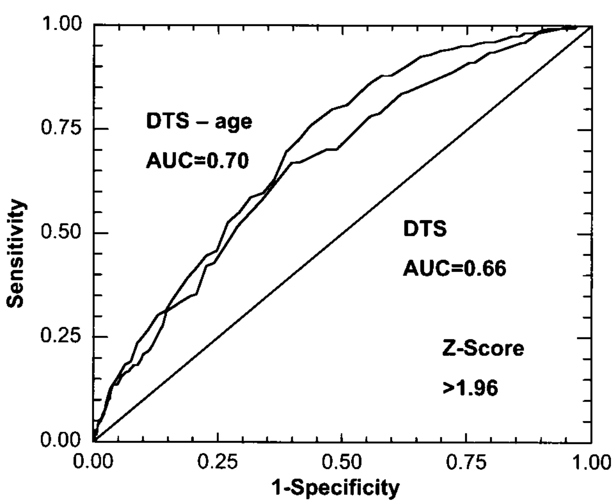

FIGURE 1. AUC of DTS vs DTS − age.

Exercise capacity was most frequently chosen in these older studies, and this has been confirmed in more recent studies. Roger et al[15] observed that workload achieved was the only exercise test variable predictive of outcomes including all-cause mortality, cardiac death, nonfatal MI, and CHF in either the young or elderly. Myers et al[16] reported that exercise capacity was a more powerful predictor of all-cause mortality than any other treadmill variable or clinical information, even when controlled for established risk factors.

Since there were a limited number of elderly included in the studies described above, it is important to consider studies including the elderly. Goraya et al[17] reported similar prognostic value of exercise capacity among elderly vs younger subjects using both overall mortality and cardiac events. Exercise capacity was associated with all-cause mortality in both age groups, and the strength of association was similar. Spin et al[18] also investigated the prognostic value of METs estimated from exercise testing among elderly using all-cause mortality.

Since the original studies of the DTS consisted largely of younger patients, it was also important to specifically evaluate it in the elderly; only two studies have done so. Unfortunately, Kwok et al[19] found that the DTS was not prognostic in patients ≥ 75 years old using cardiovascular death, MI, and cardiac interventions as the outcome. Lai et al[20] considered both death and angiographic end points and found age-specific scores to be necessary in the elderly. Given this last study, we entered the DTS and age into the Cox analysis and found them to have similar coefficients but opposite sign so that a new score equation was expressed as DTS − age. Thus, age was as strong a prognostic predictor as the DTS in our population. A score of DTS − age provided a significant improvement in AUC compared to DTS alone in the whole population and the subset of younger subjects, but there was no improvement in the elderly.

What can explain why the exercise test does not provide prognostic information in those > 75 years of age? Possibly it is due to the many competing

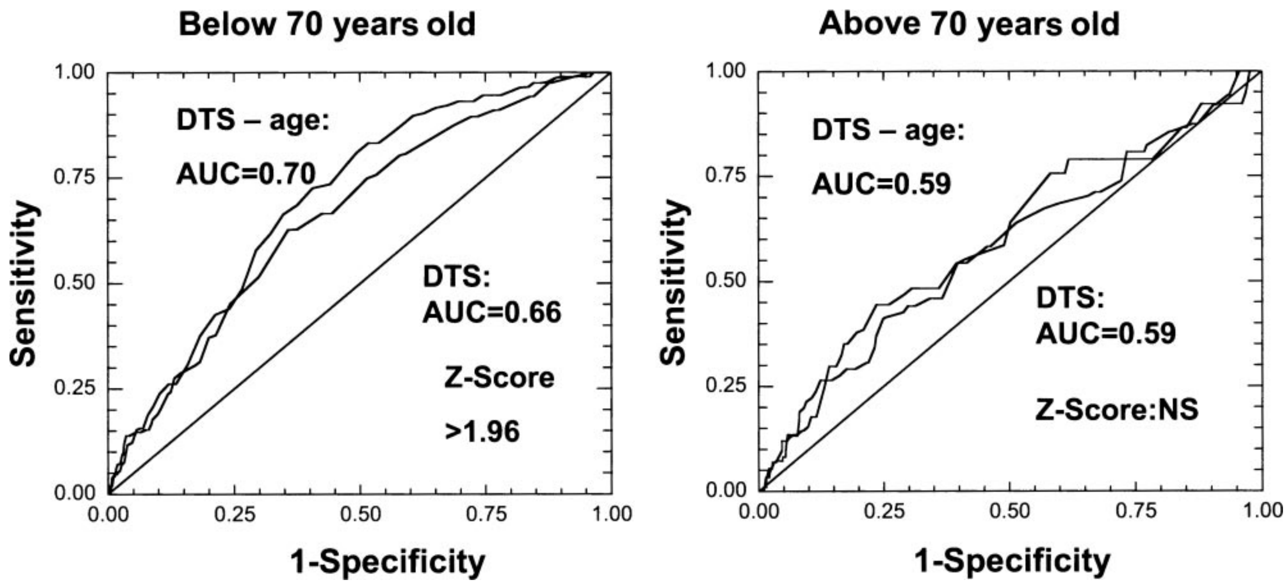

FIGURE 2. AUC of DTS and DTS − age, > 70 years and < 70 years of age. See Table 2 for expansion of abbreviation.

causes of mortality in the elderly compared to younger subjects who are more likely to die of one cause. It is also possible that the elderly are survivors who, for instance, have coronary disease but have extensive collaterals that protect them from death though not ischemia. Reduced exercise capacity in the elderly is partially explained by the high prevalence of coexisting medical problems, such as deconditioning, muscle weakness, orthopedic problems, neurologic problems, and peripheral vascular disease. Elderly patients are also more likely to have a nondiagnostic exercise ECG because of the greater prevalence of resting ECG abnormalities. These factors could confound the association between exercise test responses and outcomes.

In our study, none of the treadmill variables were selected as a predictor of outcome in those < 45 years old. This is probably due to the small number of deaths and our lack of data regarding cardiac interventions during follow-up. Exercise-induced ST-segment depression was significantly more prevalent in those who died, but it was independently associated with cardiovascular mortality only in those 45 to 55 years of age. The failure of the Duke treadmill angina score to have prognostic value in our population remains a mystery since in the very same population using the same protocol for data collection, it is one of the important predictors for the presence of angiographic disease.[21]

Our study considered a large number of patients who underwent treadmill testing for clinical indications in a general hospital/clinic setting. Patients with prior MI and/or coronary artery revascularization were excluded from the study, leaving 3,745 male veterans. Exercise testing variables were analyzed within the age groups to evaluate the effect of age and the choice of outcome: cardiovascular or all-cause death. Our results show the importance of age and the end point used in the Cox hazard analyses to develop prognostic scores. We also showed that age has a nonlinear relationship to the variables and outcomes, such that adding age to scores does not improve prediction in the elderly. This is most likely because other clinical predictors (comorbidities, psychosocial factors, and subclinical conditions) overpower the treadmill responses in the elderly even in a population such as ours in which patients with recognized heart disease were removed.

Limitations

Patients who had cardiovascular events or coronary interventions were not removed from observation since this information was not available. Data obtained from death certificates are often biased and have limited accuracy, perhaps even more so in the elderly who have competing causes of death. Lauer et al[22] pointed out the inaccuracy of death certificates in regard to attributing death to cardiovascular causes. While cardiovascular death is weaker to use than all-cause mortality in outcome studies evaluating an intervention, we believe the goal of the exercise test is to predict cardiovascular disease, not cancer, accident, suicide, or pulmonary disease. There were an unequal number of patients and clinically different selection processes for referral to

Table 6—*Exercise Test Variables Chosen in the Cox Proportional Hazards Model as Significantly and Independently Associated With Time Until Cardiovascular Death > 70 Years (n = 576) and < 70 Years (n = 3,169) of Age**

Variables	Cardiovascular Death				All-Cause Death			
	< 70 Years, Hazard Ratio (95% CI)	p Value	> 70 Years, Hazard Ratio (95% CI)	p Value	< 70 Years, Hazard Ratio (95% CI)	p Value	> 70 Years, Hazard Ratio (95% CI)	p Value
METs	0.85 (0.81–0.90)	< 0.0001		NS	0.86 (0.83–0.89)	< 0.0001	0.89 (0.83–0.95)	< 0.0005
Exercise-induced ST-segment depression	1.29 (1.11–1.49)	< 0.001	1.35 (1.09–1.66)	< 0.005		NS		NS
Maximal heart rate		NS		NS	0.99 (0.99–1.00)	< 0.05		NS

*See Table 2 for expansion of abbreviations.

treadmill testing in each age group. We lacked information regarding comorbidities other than pulmonary and cardiac diseases and have no data on heart rate recovery in this population. Certainly applying nonlinear statistical models could provide a general equation, but a score developed in such a way would simply lessen the effect of exercise test responses relative to age in the elderly and offer no real advantage to clinicians.

Conclusion

Both age and the outcome selected as an end point affect the exercise test responses chosen for scores to predict prognosis. Differences in either age of the subjects tested and the outcome selected as the end point can explain the differences in the studies using exercise testing to predict prognosis. The possibility that the exercise test does not have prognostic value in men > 75 years of age has such clinical importance that research using a prospective design avoiding the limitations of our retrospective analysis should be accomplished as soon as possible.

References

1 Froelicher VF, Myers J. Exercise and the Heart. 4th ed. Philadelphia, PA: Saunders-Mosby, 2000; 222–224

2 Ustin J, Umann T, Froelicher V. Data management: a better approach. Physicians Comput 2000; 12:30–33

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

4 Wolthuis R, Froelicher VF, Fischer J, et al. New practical treadmill protocol for clinical use. Am J Cardiol 1977; 39:697–700

5 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

6 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

7 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

8 Reid M, Lachs M, Feinstein A. Use of methodological standards in diagnostic test research. JAMA 1995; 274:645–651

9 Yang JC, Wesley RC, Froelicher VF. Ventricular tachycardia during routine treadmill testing: risk and prognosis. Arch Intern Med 1991; 151:349–353

10 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

11 Mark DB, Hlatky MA, Harrell FE, et al. Exercise treadmill score for predicting prognosis in coronary artery disease. Ann Intern Med 1987; 106:793–800

12 Mark DB, Shaw L, Harrell FE Jr, et al. Prognostic value of a treadmill exercise score in outpatients with suspected coro-

nary artery disease. N Engl J Med 1991; 325:849–853
13 Gibbons RJ, Balady GJ, Timothy Bricker J, et al. ACC/AHA 2002 guideline update for exercise testing. J Am Coll Cardiol 2002; 40:1531–1540
14 Lissin LW, Gauri AJ, Froelicher VF, et al. The prognostic value of body mass index and standard exercise testing in male veterans with congestive heart failure. J Card Fail 2002; 8:206–215
15 Roger VL, Jacobsen SJ, Pellikka PA, et al. Prognostic value of treadmill exercise testing: a population-based study in Olmsted County, Minnesota. Circulation 1998; 98:2836–2841
16 Myers J, Prakash M, Froelicher VF, et al. Exercise capacity and mortality among men referred for exercise testing. N Engl J Med 2002; 346:793–801
17 Goraya TY, Jacobsen SJ, Pellikka PA, et al. Prognostic value of treadmill exercise testing in elderly persons. Ann Intern Med 2000; 132:862–870
18 Spin JM, Prakash M, Froelicher VF, et al. The prognostic value of exercise testing in elderly men. Am J Med 2002; 112:453–459
19 Kwok JM, Miller TD, Hodge DO, et al. Prognostic value of the Duke treadmill score in the elderly. J Am Coll Cardiol 2002; 39:1475–1481
20 Lai S, Kaykha A, Yamazaki T, et al. Treadmill scores in elderly men. J Am Coll Cardiol 2004; 43:100–107
21 Raxwal V, Shetler K, Morise A, et al. Simple treadmill score to diagnose coronary disease. Chest 2001; 119:1933–1940
22 Lauer MS, Blackstone EH, Young JB, et al. Cause of death in clinical research. J Am Coll Cardiol 1999; 34:618–620