Journal of the American College of Cardiology
© 2004 by the American College of Cardiology Foundation
Published by Elsevier Inc.

Vol. 43, No. 4, 2004
ISSN 0735-1097/04/$30.00
doi:10.1016/j.jacc.2003.07.051

Treadmill Scores in Elderly Men

Steve Lai, MD,† Amir Kaykha, MD,* Takuya Yamazaki, MD,* Mary Goldstein, MD,†
Joshua M. Spin, MD, PHD,* Jonathan Myers, PHD,* Victor F. Froelicher, MD†
Stanford and Palo Alto, California

OBJECTIVES This study seeks to further characterize the role of exercise testing in the elderly for prognosis and diagnosis of coronary artery disease.

BACKGROUND Recent exercise testing guidelines have recognized that statements regarding the elderly do not have an adequate evidence-based quality because the studies they are based on have limitations in sample size and design. The Duke Treadmill Score has been recommended for risk stratification, but recent evidence has suggested that it does not function in the elderly.

METHODS The study population consisted of male veterans (1,872 patients ≥65 years; 3,798 patients <65 years) who underwent routine clinical exercise testing with a mean follow-up of six years. A subset who underwent coronary angiography as clinically indicated (elderly, n = 405; younger, n= 809) were included. The primary outcome for all subjects was cardiovascular mortality with coronary angiographic findings as the outcome in those selected for angiography.

RESULTS In survival analysis, exercise-induced ST depression was prognostic in both age groups only when cardiovascular death was considered as the end point. Peak metabolic equivalents were the most significant predictor for both age groups only when all-cause death was considered as the end point. New age-specific prognostic scores were developed and found to be predictive for cardiovascular mortality in the elderly. Moreover, in the angiographic subset of the elderly, a specific diagnostic score provided significantly better discrimination than exercise ST measurements alone. For any new score, there is a need for validation in another elderly population.

CONCLUSIONS The mortality end point affected the choice of prognostic variables. This study demonstrates that exercise test scores can be helpful for the diagnosis and prognosis of coronary disease in the elderly. (J Am Coll Cardiol 2004;43:606–15) © 2004 by the American College of Cardiology Foundation

With aging of the population, heart disease continues to be the leading cause of mortality in the U.S. The standard exercise test is widely used to evaluate heart disease in the elderly. However, the value of exercise testing has not been fully characterized for elderly patients. Elderly persons exhibit a greater prevalence and severity of coronary disease and comorbid conditions, which may alter the ability of exercise testing to predict outcomes (1). In addition, differences in physiological responses to exercise may appear with aging (2–4). A recent study by Kwok et al. (5) concluded that the widely recommended Duke Treadmill Score (DTS) was limited in predicting cardiovascular mortality in the elderly.

In accordance with exercise testing guideline recommendations (6,7), we sought to obtain more data by examining the utility of exercise testing in male veterans and the DTS. We recently confirmed (8) a prior study of 514 elderly patients that found that exercise-induced ST depression did not independently predict prognosis in the elderly (9). However, the two studies (8,9) only considered all-cause mortality and did not include angiographic data. In this analysis, a larger population is considered, including those with established heart disease, and both angiographic results and cardiovascular death are included as outcomes.

METHODS

Population. After exclusion of patients with a history and/or findings of congestive heart failure, use of digoxin, or more than 1 mm of resting ST depression, the population consisted of 5,625 consecutive patients referred to two clinical exercise laboratories (Long Beach Veterans Affairs 1987 to 1991; Palo Alto Veterans Affairs 1992 to 2000). These labs were directed in consistent fashion by two of the authors (V.F.F. and J.M.), and patients already enrolled in research protocols were not considered in the analyses. This study included a subset of the population free of established coronary disease (prior myocardial infarction [MI] or coronary intervention) that had been selected for clinical indications for coronary angiography (≥65 years, n = 405; ≤65 years, n = 809) within four months of testing. The study protocol was approved by the hospital review committee at each Veterans Affairs institution. Written informed consent was obtained from all patients.

Data collection. Both labs 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 read by two of the investigators (V.F.F. and J.M.). No imaging modality was performed in conjunction with exercise testing. A

From the *Division of Cardiovascular Medicine, Stanford University Medical Center, Stanford, California; and the †Veterans Affairs Health Care System, Palo Alto, California. Views expressed are those of the authors and do not necessarily reflect those of the Department of Veterans Affairs.

Manuscript received March 13, 2003; revised manuscript received July 2, 2003, accepted July 15, 2003.

Abbreviations and Acronyms

AUC	=	area under the curve
CABG	=	coronary artery bypass graft
CAD	=	coronary artery disease
DTS	=	Duke Treadmill Score
LAD	=	left anterior descending coronary artery
MET	=	metabolic equivalent
MI	=	myocardial infarction
PRP	=	pressure rate product
ROC	=	receiver operator characteristic
VA/UWV	=	Veterans Affairs/University of West Virginia angiographic diagnostic score

thorough clinical history, a list of medications, and cardiac risk factors were recorded prospectively at the time of testing using computerized forms (10,11). The forms included standard definitions of clinical conditions and exercise responses.

Exercise testing. Patients underwent symptom-limited treadmill testing using the United States Air Force School of Aerospace Medicine (USAFSAM) protocol (12) or an individualized ramp treadmill protocol (13,14). Heart rate targets were not used as an end point or to judge the adequacy of the test. Patients were placed supine immediately after exercise (15). No medications were changed or stopped before testing, and no test was classified as indeterminate (16). After maximal heart rate was regressed on age in our male veteran population, the linear regression equation of $210 - 0.8 \times$ (age) was derived for age-predicted maximal heart rate. Trained physicians or nurses were always in direct attendance of the test, and the senior authors were always available for consultation.

Visual ST-segment depression was measured at the J junction and corrected for pre-exercise ST-segment depression while standing; ST slope was measured over the following 60 ms and classified as upsloping, horizontal, or downsloping. The ST 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 or more 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 (17). Blood pressure was taken manually. Estimated metabolic equivalents (METs) were calculated from treadmill speed and grade. The same equation used by the DTS group was applied to convert from Bruce minutes to METs: (Bruce minutes) = (METs + 2.2)/1.3. Pressure rate product (PRP or double product) pre-test and at maximal exercise were calculated as the product of systolic blood pressure and heart rate. Delta PRP was calculated as follows: PRP at maximal exercise minus PRP at rest/1,000. A previously validated treadmill angiographic score was calculated in each patient (18).

Coronary angiography. Coronary artery narrowing was visually estimated and expressed as percent lumen diameter stenosis. Patients with a 50% narrowing in one or more of the following were considered to have significant angiographic coronary artery disease (CAD): the left main coronary artery, the left anterior descending artery (LAD), the left circumflex, the right coronary artery, or major branches of these vessels. Severe disease consisted of either three-vessel disease, or two vessels with 50% narrowing or greater (requiring inclusion of either the proximal LAD or the left main). The 50% lesion criterion was chosen to be consistent with the cooperative trialists choice (19), and because we found little discriminatory difference between 50% to 75% lesions (20). Decisions regarding the need for cardiac catheterization were consistent with clinical practice (i.e., no attempt was made to limit work up bias).

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. Mean follow-up was six years. Data on subsequent interventions or nonfatal cardiovascular events were not available. Death status was determined as of July 2000. Cardiovascular death was differentiated from all-cause death in the California death registry by death certificates.

Statistical methods. The population was divided into patients ≥65 years of age (elderly) and those <65 years of age. Comparisons were performed between these groups, and between survivors and nonsurvivors within each age group. Number Cruncher Statistical Systems Software (Salt Lake City, Utah) was used for all statistical analyses after transferring the data from an ACCESS (Microsoft) database. All relevant clinical and the exercise variables were selected for analysis. Unpaired *t* testing was used to compare mean values for historical and demographic electrocardiogram (ECG) and exercise test data between those in the younger versus elderly group and for survivors versus nonsurvivors. Chi-square tests were used for differences between proportions. All-cause and cardiovascular mortality were used as the end point for follow-up for survival analysis. Patients were not removed from observation for interventions or infarction because this information was not available. Survival analysis was performed using Kaplan-Meier curves to compare variables and cut points. Cox hazard function was used for survival analysis to demonstrate which variables were independently associated with survival.

The scores were compared to the DTS by the use of the Cox hazard function, receiver operating characteristic (ROC) curves analysis, and Kaplan-Meier survival curves. For the Kaplan-Meier comparisons, the populations were divided into tertiles such that one-third of the population was in each score group associated with low, intermediate, and high risk.

In the subset clinically chosen for angiography, subjects were separated into those with and without significant angiographic disease. How well ST-segment depression and the scores separated patients with and without a given outcome (CAD) was assessed in the age groups by means of standard 2×2 tables and the area under ROC curve. A Z-score was calculated to determine if the differences

JACC Vol. 43, No. 4, 2004
February 18, 2004:606–15

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

Choose only one per group

Figure 1. The Veterans Affairs/University of West Virginia score for diagnosing angiographic coronary artery disease.

between the areas under the curve (AUCs) were statistically significant (Z > 1.96 = p value <0.001).

Score derivation. The DTS was calculated as exercise time (converted to Bruce minutes from METs) − 5 × (amount ST depression) − 4 × (treadmill angina index). Variables found to be univariately significantly different between those who died and those who survived were considered in the Cox model. Variables then found to be significantly and independently associated with time to cardiovascular death were selected for predictive scores. Their coefficients were divided by the smallest coefficient to determine a relative weight with the weight of the smallest variable being unity. These weights were then multiplied by the variable and all products summed to provide new prognostic scores specific for the age groups. The previously validated simplified Veterans Affairs/University of West Virginia angiographic diagnostic score (VA/UWV score) (Fig. 1) was calculated as 6 × maximal heart rate code + 5 × ST depression code + 4 × age code + angina pectoris code + hypercholesterolemia code + diabetes code + treadmill angina index code (18). The treadmill angina index has a value of 0 if the patient had no angina during exercise, 1 if the patient had non-limiting angina, and 2 if angina was the reason the patient stopped exercising.

RESULTS

Population demographics. Average resting heart rate was ~3 beats/min lower in the elderly, with a corresponding mean systolic blood pressure 7 mm Hg higher. Elderly subjects had a mean body mass index 1.6 kg/m^2 lower than the younger subjects. There were no significant differences in age between the different ethnicities. Overall, 74% were Caucasian, 9% were Hispanic, and 12% were African American. The relevant variables for elderly and younger patients are in Table 1. There were no significant differences in beta-blocker (19%, elderly; 19%, younger), anticoagulant, or antiarrhythmic usage between the age groups, but the elderly were more commonly on other cardiovascular medications, such as calcium channel blockers and nitrates. The elderly were significantly more likely to have a history of pulmonary disease, atrial fibrillation, claudication, and stroke. Elderly patients were also more likely to have a history of MI and coronary interventions such as coronary artery bypass graft surgery (CABG) or percutaneous coronary intervention. The only cardiac risk factor more common in the elderly was hypertension (55% vs. 44%). In fact, smoking (50% vs. 60%) and obesity (44% vs. 57%) were less prevalent in the elderly. Diabetes (11% vs. 10%) and

Table 1. Demographic and Clinical Characteristics, With Univariate Comparisons Between Ages Above and Below 65 Years for Total, Surviving, and Nonsurviving Patients From Cardiac Causes of Death

	Total Population (n = 5,625)			Age <65 (n = 3,798)			Age ≥65 (n = 1,827)		
	Age <65	Age ≥65	p Value	Survived	CV Death	p Value	Survived	CV Death	p Value
Demographics, n (%)	3,798 (68)	1,827 (32)		3,560 (94)	238 (6)		1,593 (87)	234 (13)	
Age, yrs	52.8 ± 8.4	70.8 ± 4.6	< 0.001	52.5 ± 8.5	57.1 ± 6.2	< 0.001	70.7 ± 4.6	71 ± 4.6	0.39
Medication use									
Nitrates, n (%)	787 (20.7)	460 (25.2)	< 0.001	686 (19.3)	101 (42.4)	< 0.001	360 (22.6)	100 (42.7)	< 0.001
Antihypertensives, n (%)	720 (19)	535 (29.3)	< 0.001	640 (18)	80 (33.6)	< 0.001	442 (27.7)	93 (39.7)	< 0.001
Medical history and cardiac risk factors claudication, n (%)	160 (4.2)	118 (6.5)	< 0.001	142 (4)	18 (7.6)	0.01	91 (5.7)	27 (11.5)	0.001
Atrial fibrillation, n (%)	42 (1.1)	38 (2.1)	0.004	41 (1.2)	1 (0.4)	0.37	31 (1.9)	7 (3)	0.32
Stroke, n (%)	96 (2.5)	94 (5.1)	< 0.001	85 (2.4)	11 (4.6)	0.05	80 (5)	14 (6)	0.53
Pulmonary disease, n (%)	193 (5.1)	162 (8.9)	< 0.001	174 (4.9)	19 (8)	0.05	141 (8.9)	21 (9)	0.90
Typical angina, n (%)	1,039 (27.4)	568 (31.1)	0.004	939 (26.4)	100 (42)	< 0.001	458 (28.8)	110 (47)	< 0.001
MI history/Q waves, n (%)	967 (25.5)	569 (31.1)	< 0.001	837 (23.5)	130 (54.6)	< 0.001	452 (28.4)	117 (50)	< 0.001
CABG, n (%)	236 (6.2)	226 (12.4)	< 0.001	193 (5.4)	43 (18.1)	< 0.001	169 (10.6)	57 (24.4)	< 0.001
PCI, n (%)	233 (6.1)	108 (5.9)	0.77	216 (6.1)	17 (7.1)	0.49	95 (6)	13 (5.6)	0.88
Hypertension, n (%)	1,651 (43.5)	1,008 (55.2)	< 0.001	1,526 (42.9)	125 (52.5)	0.004	866 (54.4)	142 (60.7)	0.08

CABG = coronary artery bypass graft surgery; CV = cardiovascular; MI = myocardial infarction; PCI = percutaneous coronary intervention.

hypercholesterolemia (29% vs. 31%) did not differ significantly between the age groups. Resting ECG abnormalities, including Q waves, atrial fibrillation, prolonged QRS duration, ST depression, and left ventricular hypertrophy were more common in the elderly (Table 2).

Exercise test responses. No major complications were encountered during testing, although there were six incidents of nonsustained ventricular tachycardia. Exercise test responses for the elderly and the younger veterans are presented in Table 2. Angina occurred during testing in 12% of the population, and it was the reason for stopping in 6%. All of the major exercise ECG abnormalities were significantly more prevalent in the elderly. Abnormal exercise-induced ST depression, the most common ECG abnormality, was more prevalent in the elderly (31.5% vs. 19.2%, $p < 0.001$). Silent ischemia (i.e., ST depression without chest pain) was nearly twice as common in the elderly. The next most prevalent ECG abnormality was frequent premature ventricular contractions and/or ventricular tachycardia (three beats in a row or more), and this combined response was twice as common in the elderly (11% vs. 5%, $p < 0.001$). ST elevation was rare (<1.5%). Both age groups gave a similar effort as reflected in the mean Borg scale rating of 17 for "very hard."

Average maximal heart rate and METs achieved were significantly lower in the elderly. Failure to reach age-predicted heart rate target occurred in almost half of those tested, and was slightly, but significantly, more common in the elderly. The maximal PRP and delta PRP were significantly lower in the elderly.

Survival analysis. The elderly age group experienced 28.5% all-cause mortality (44.9% cardiac deaths) versus 14.3% all-cause mortality (43.8% cardiac deaths) in the younger age group. The average annual cardiovascular mortality was 1.4% in the elderly and 0.7% in the younger veterans. Nonsurvivors in both age groups were significantly more likely to be taking cardiovascular medications. Also, patients with a history of MI or previous coronary interventions were less likely to survive, regardless of age group. Resting ECG abnormalities were associated with increased cardiovascular mortality in all groups. ST depression during exercise was significantly more prevalent in those who died compared with those who survived in both age groups (33% vs. 18%, younger age group; 45% vs. 30%, elderly). All measured exercise hemodynamic variables were significantly different, that is, lower in nonsurvivors versus survivors in both age groups.

Using stepwise selection, the proportional hazards model was allowed to build on each variable group. The clinical variables and exercise variables were considered univariately, and those that were significant ($p < 0.001$) were analyzed multivariately. Table 3 shows the significant clinical and exercise variables that were independently associated with cardiovascular and all-cause mortality. In the elderly group, resting ST depression, delta PRP, history of CABG, MI history/Q waves, and abnormal ST depression during exercise were the most significant clinical and exercise test variables associated with cardiovascular death. In the younger age group, the maximal METs, history of CABG, prior history of MI/Q waves, and abnormal ST depression during exercise were most significant. For all-cause mortality analysis, the most important exercise test variable was maximal METs in both age groups, and exercise-induced ST depression was no longer chosen. Anti-ischemic medications such as calcium channel blockers and beta-blockers can produce false negative tests on treadmill testing. After adjusting for calcium channel blockers and beta-blockers, there were no differences in the clinical and exercise test variables in the hazard model. An age-specific prognostic score was created for patients ≥65 years as: [Δ PRP (10^{-3}) − 10 (rest ST depression yes/no + history of MI or Q-wave yes/no + CABG yes/no) − 4 (exercise ST depression in mm)]. For patients <65 years, the score was: [peak METs − 5 (CABG yes/no + history of MI or Q-wave yes/no) − 2 (exercise ST depression in mm)].

Age-specific prognostic scores compared with the Duke score. Table 4 shows the comparison of AUC for our >65 and <65 age-specific scores versus the DTS. Comparing our scores with the DTS for the total population, the ROC scores were significantly different and demonstrated the increased discriminatory power of the specific age scores in both age groups. For the >65 group, the AUC (± standard error) of the ROC plot for the DTS and Veterans Administration score was 0.59 ± 0.008 and 0.69 ± 0.007 (Fig. 2). For the <65 group, the AUC of the ROC plot for the DTS and Veterans Administration score was 0.67 ± 0.007 and 0.73 ± 0.006. When patients with previous MI and CABG/percutaneous coronary intervention were excluded, the Veterans Administration and DTS ROC curves were still significantly different in the >65 years group (0.65 ± 0.011 vs. 0.60 ± 0.012). However, the ROC curves for the Veterans Administration and DTS scores were not significantly different in the <65 year group (Fig. 3). In Figure 4, we show Kaplan-Meier survival curves of DTS using original cut points and the age-specific score for >65 using tertile cut points. In our age-specific score, the low-risk and high-risk group had more separation than that of the DTS.

Diagnostic characteristics of the scores in those with angiograms. In the angiographic subset (elderly, n = 405; younger, n = 809), the prevalence of angiographic disease was significantly higher in the elderly (72% vs. 53%, $p < 0.001$). Patients with CAD in both age groups had a significantly higher prevalence of hypercholesterolemia, typical angina, and abnormal exercise testing. Patients <65 years with CAD had ~1.7 MET lower exercise tolerance than those without CAD ($p < 0.001$). Of those <65, 33% had abnormal exercise tests, and in those >65, 49% had abnormal exercise tests compared with 21% and 33%, respectively, in the total population consistent with workup bias.

There were no significant differences in test characteristics for the standard criterion of 1 mm of ST depression

Table 2. Resting and Exercise Test Values, With Univariate Comparison Between Ages Above and Below 65 Years for Total, Surviving, and Nonsurviving Patients From Cardiac Causes of Death

	Total Population (n = 5,625)			Age <65 (n = 3,798)			Age ≥65 (n = 1,827)		
	Age <65	Age ≥65	p Value	Survived	CV Death	p Value	Survived	CV Death	p Value
Resting value, n (%)	3,798 (68)	1,827 (32)		3,560 (94)	238 (6)		1,593 (87)	234 (13)	
Abnormal ECG, n (%)	1,247 (32.8)	841 (46)	< 0.001	1,116 (31.3)	131 (55)	< 0.001	705 (44.3)	136 (58.1)	< 0.001
RBBB, LBBB, IVCD, n (%)	252 (6.6)	246 (13.5)	< 0.001	223 (6.3)	29 (12.2)	< 0.001	205 (12.9)	41 (17.5)	0.06
LVH, n (%)	156 (4.1)	98 (5.4)	0.03	137 (3.8)	19 (8)	0.004	75 (4.7)	23 (9.8)	0.002
Atrial fibrillation, n (%)	17 (0.4)	31 (1.7)	< 0.001	14 (0.4)	3 (1.3)	0.09	26 (1.6)	5 (2.1)	0.58
ST depression, n (%) (mean in mm)	240 (6.3) (0.4)	208 (11.4) (0.4)	< 0.001	206 (5.8)	34 (14.3)	< 0.001	155 (9.7)	53 (22.6)	< 0.001
Q waves, n (%)	595 (15.7)	362 (19.8)	< 0.001	518 (14.6)	77 (32.4)	< 0.001	290 (18.2)	72 (30.8)	< 0.001
Heart rate, mean ± SD	78.2 ± 15.2	75.1 ± 14.6	< 0.001	78.2 ± 15.1	78.4 ± 16.7	0.88	75 ± 14.6	75.8 ± 14.8	0.46
Systolic BP, mean ± SD	131.4 ± 20.2	138.7 ± 21.1	< 0.001	131.4 ± 20.0	130.5 ± 22.5	0.5	138.8 ± 20.9	138.2 ± 22.8	0.72
Exercise value									
Exercise-induced ST depression, n (%)	728 (19.2)	575 (31.5)	< 0.001	649 (18.2)	79 (33.2)	< 0.001	470 (29.5)	105 (44.9)	< 0.001
Silent ischemia, n (%)	469 (12.3)	370 (20.3)	< 0.001	425 (11.9)	44 (18.5)	0.004	307 (19.3)	63 (26.9)	< 0.007
Rate-dependent LBBB, RBBB, IVCD, n (%)	120 (3.2)	115 (6.3)	< 0.001	106 (3)	14 (5.9)	0.02	96 (6)	19 (8.1)	0.25
Exercise-induced arrythmias VT or frequent PVCs, n (%)	198 (5.2)	208 (11.4)	< 0.001	176 (4.9)	22 (9.2)	0.006	167 (10.5)	41 (17.5)	0.002
Maximal heart rate, mean ± SD	141.6 ± 24.6	128.2 ± 22.4	< 0.001	142.7 ± 24.2	125.1 ± 24.9	< 0.001	128.9 ± 22.2	122.9 ± 23.4	< 0.001
Exercise heart rate <85 age predicted, n (%)	1,858 (48.9)	985 (53.9)	< 0.001	1,690 (47.5)	168 (70.6)	< 0.001	833 (52.3)	152 (65.0)	< 0.001
Maximal PRP, mean ± SD (× 1,000)	25.6 ± 6.9	23.2 ± 6.2	< 0.001	25.9 ± 6.8	21.6 ± 7.2	< 0.001	23.5 ± 6.2	21.3 ± 5.9	< 0.001
Delta PRP ± SD (× 1,000)	15.4 ± 6.3	12.8 ± 5.6	< 0.001	15.6 ± 6.2	11.4 ± 6.1	< 0.001	13.1 ± 5.7	10.8 ± 5	< 0.001
METs, mean ± SD	9.1 ± 3.7	6.6 ± 2.9	< 0.001	9.3 ± 3.7	6.9 ± 3.1	< 0.001	6.7 ± 2.9	6.0 ± 2.8	0.002

BP = blood pressure; CV = cardiovascular; ECG = electrocardiogram; IVCD = intraventricular conduction defect; LBBB = left bundle branch block; LVH = left ventricular hypertrophy; MET = metabolic equivalent; PRP = pressure rate product; PVC = premature ventricle contraction; RBBB = right bundle branch block; VT = ventricular tachycardia.

Table 3. Clinical and 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

Cause of Death	Age	Variable	Regression Coefficient	Hazard Ratio	95% CI	p Value
Cardiac	≥65	Resting ST depression	0.53	1.70	1.22–2.36	< 0.001
		Delta PRP	−0.04	0.96	0.93–0.99	< 0.001
		CABG history	0.56	1.74	1.27–2.40	< 0.001
		MI history/Q waves	0.43	1.53	1.22–1.93	0.003
		Abnormal ST during exercise	0.17	1.18	1.05–1.33	0.004
		Hypertension	0.33	1.39	1.06–1.83	0.006
		METs	−0.07	0.93	0.87–0.98	0.04
Cardiac	<65	METs	−0.12	0.89	0.84–0.93	< 0.001
		CABG history	0.64	1.90	1.34–2.68	< 0.001
		MI history/Q waves	0.55	1.73	1.31–2.30	< 0.001
		Abnormal ST during exercise	0.21	1.23	1.09–1.40	0.002
		Hypertension	0.34	1.40	1.08–1.82	0.01
		Maximal heart rate	−0.01	0.99	0.98–1.00	0.02
All causes	≥65	METs	−0.09	0.92	0.88–0.95	< 0.001
		Resting ST depression	0.41	1.50	1.18–1.90	< 0.001
		Delta PRP	−0.04	0.96	0.94–0.98	0.002
		Resting heart rate	0.007	1.01	1.00–1.01	0.02
All causes	<65	METs	−0.09	0.91	0.88–0.94	< 0.001
		Abnormal resting ECG	0.33	1.40	1.17–1.67	< 0.001
		CABG history	0.34	1.41	1.08–1.84	0.007
		Resting heart rate	0.015	1.01	1.01–1.02	0.01
		Maximal heart rate	−0.011	0.99	0.98–0.99	0.006
		Current smoker	0.19	1.21	1.01–1.44	0.04

CABG = coronary artery bypass grafting; CI = confidence interval; ECG = electrocardiogram; MET = metabolic equivalent; MI = myocardial infarction; PRP = pressure rate product.

(predictive accuracy of 59% for the elderly and 65% for the younger group, sensitivity of 55% for the elderly and 47% for the younger group). The AUC of the ROC curves for ST depression, DTS, and a previously validated diagnostic score (VA/UWV score) (Fig. 1) (18) are presented in Table 5 for the two age groups. The Z-score was calculated to compare the ability to discriminate between the age groups and then for the scores compared with the ST measurements alone. For the younger group, the AUC (± standard error) of the ROC plot for the ST response alone, DTS, and VA/UWV score were 0.67 ± 0.019, 0.72 ± 0.018, and 0.79 ± 0.016, respectively. For the elderly population, the AUC of the ROC plot for the ST response alone, DTS, and VA/UWV score were 0.66 ± 0.029, 0.72 ± 0.026, and 0.75 ± 0.025, respectively, and are shown in Figure 5. These were not significantly different between the age groups. In patients <65 years, AUC for VA/UWV score was significantly greater than the ST response alone and DTS, but both scores were significantly better than the ST measurements alone. For the elderly, only the AUC for VA/UWV score was significantly greater than that of ST response alone.

DISCUSSION

Background. 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 (21,22). A recent study by Kwok et al. (5) comparing 247 patients 75 years or older to younger patients demonstrated that the prognostic value of the DTS in the elderly was limited. This could be explained by the fact that they have a higher prevalence of CAD, more

Table 4. Comparison of AUC for Above and Below 65 Years Age-Specific Prognostic Scores and the Duke Treadmill Score

Population	Age	Score	Median	5%	95%	Mean ± SD	AUC ± SE	Z-Score§
Total	≥65	DTS*	4.0	−11.9	11.0	2.3 ± 7.0	0.59 ± 0.008	
		Above 65 score†	6.1	−16.6	20.9	4.7 ± 11.5	0.69 ± 0.007	9.35
Total	<65	DTS	7.4	−7.1	15.0	6.4 ± 6.8	0.67 ± 0.007	
		Below 65 score‡	7.1	−3.0	15.0	6.8 ± 5.6	0.73 ± 0.006	7.09
MI & CABG excluded	≥65	DTS	4.5	−11.5	11.4	3.0 ± 6.9	0.60 ± 0.012	
		Above 65 score	11.1	−4.2	22.1	10.5 ± 8.0	0.65 ± 0.011	3.09
MI & CABG excluded	<65	DTS	8.1	−5.5	15.2	7.4 ± 6.5	0.65 ± 0.011	
		Below 65 score	9.0	2.0	15.4	9.0 ± 4.2	0.67 ± 0.010	1.02 (NS)

*DTS = Duke Treadmill Score = METs −5 exercise ST depression −4 (Duke treadmill angina pectoris score); †Above 65 score = Δ PRP (10^{-3}) −10 (rest ST depression yes/no + MIhxQwave yes/no + CABG yes/no) −4 (exercise ST depression); ‡Below 65 score = METs −5 (CABG yes/no + MIhxQwave yes/no) −2 (exercise ST depression); §Z-score > 1.96 = ROC curves are significantly different (note that the means and medians are provided to illustrate when the scores are non-Guassian).

AUC = area under the curve; CABG = coronary artery bypass graft surgery; MI = myocardial infarction; NS = no significant difference; PRP = pressure rate product; ROC = receiver operator characteristic.

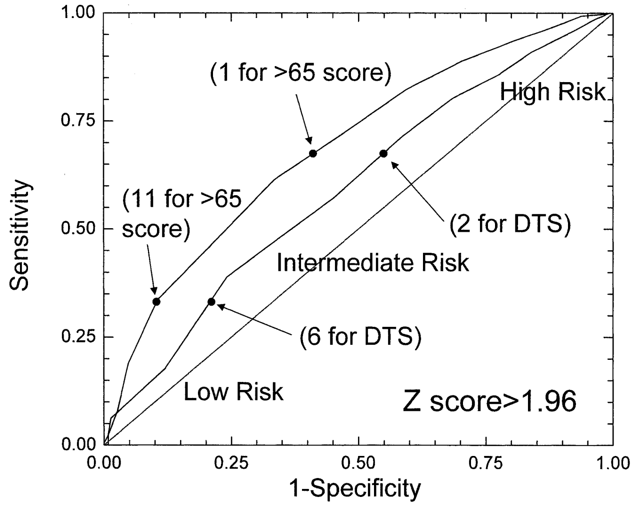

Figure 2. Receiver operator characteristic plot for the age-specific and Duke Treadmill Scores (DTS) in patients ≥65 years of age with the end point being cardiovascular death.

severe CAD, and/or a lower exercise capacity. We included patients with minor resting ST depression (<1 mm) consistent with the guidelines and with the following two studies. Kwok et al. (23) validated the prognostic accuracy of the DTS in patients with nonspecific ST-depression, and Fearon et al. (24) also demonstrated that minor degrees of resting ST-segment depression did not affect the diagnostic characteristics of the treadmill test. Both demonstrated that those with resting ST depression have a higher mortality, making them important to evaluate further. While the DTS was designed to provide survival estimates, Shaw et al. (25) verified its diagnostic usefulness for angiographic CAD. However, in a similar mixed-age population with angiographic CAD as the outcome, we compared the diagnostic characteristics of the ST response alone, the DTS, and three

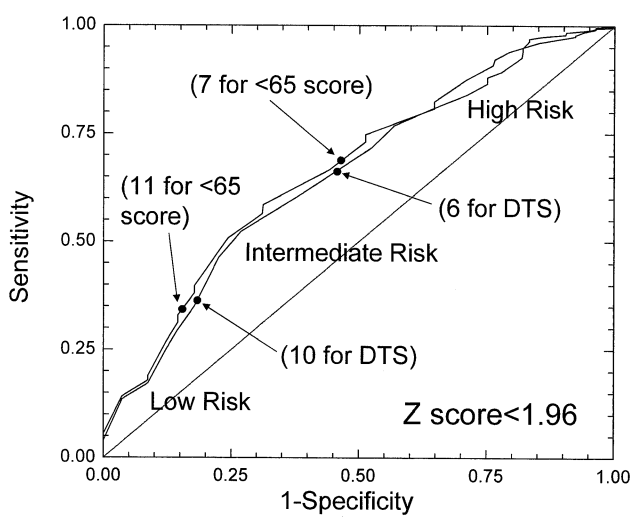

Figure 3. Receiver operator characteristic plot for the age-specific and Duke Treadmill Scores (DTS) in patients <65 years of age with patients with known coronary artery disease excluded and the end point being cardiovascular death.

other scores and found that the DTS was less accurate than other treadmill scores (26).

Our study showed that the mortality end point affected the choice of prognostic variables. In addition, the above 65 age group had different clinical and exercise variables of prognostic utility. The DTS has previously been validated for prognostication. The DTS was developed in a referral population that was younger (mean age, 49 years) (27). In addition, the DTS did not incorporate cardiac risk factors and comorbid conditions that we used in our multivariate analysis. The Veterans Administration is the largest health care system in the world with an older population, and, if the DTS is not appropriate for the elderly, there must be a search for a more predictive score. In this Veterans Administration cohort, a new prognostic score for age above 65 is proposed. This is adequate to provide an initial evaluation of the score. However, validation in another geographically distinct group is required to assess the transportability or generalizability.

The differences in the hemodynamic responses likely reflect physiological changes with aging, and their validity is supported by the similarly perceived exertion in both groups. The ability of the coronary circulation to respond to myocardial metabolic demand is estimated by the product of systolic blood pressure and heart rate, generally called PRP or double product (28,29). Generally, peak exercise systolic and delta systolic blood pressure increase with advancing age while PRP decreases, similar to our results (30). Prognostic studies demonstrating the prognostic value of delta PRP have largely centered on post-MI patients (31,32). The only prior study to consider and find PRP significantly associated with time to death in non-MI patients was from the Seattle Veterans Administration (33).

Comparison with younger patients. As expected, the elderly group experienced a higher cardiovascular mortality. The notable finding in the elderly group was that resting ST depression and delta PRP were the most significant exercise test variables associated with cardiovascular death, as opposed to maximal METs being the most powerful in the younger population. It is important to note that maximal METs had similar risk ratios in both groups, but, in the elderly, other clinical and exercise variables became more significant. Table 2 indicates the standard deviation of METs is smaller for the elderly when compared with the younger group. This is expected as the METs achieved decreases with age. However, we have previously found that the percent age-adjusted METs did not function better than absolute METs achieved (32).

Resting ST depression is more prevalent in the elderly, and our findings are consistent with previous studies that revealed that patients with resting ST abnormalities have a worse prognosis (23,24). It is interesting to note that our derived score for the younger group looked very similar to the DTS when used in patients without CAD. The Duke treadmill angina index was evaluated as a potential predictor in all of our

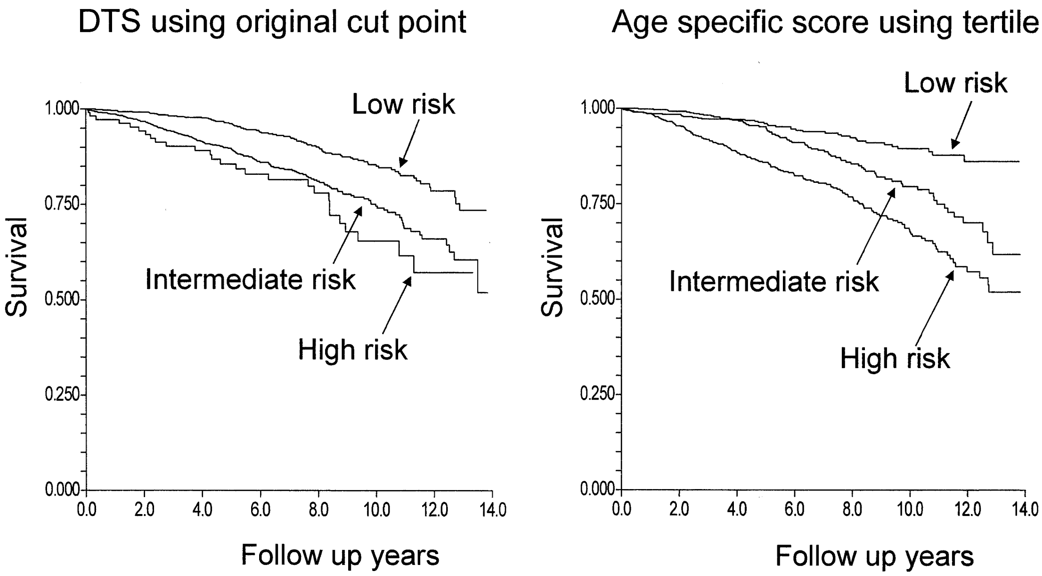

Figure 4. Kaplan-Meier survival curves of Duke Treadmill Scores (DTS) and age-specific prognostic score.

prognostic models and was not chosen as significantly and independently associated with time to death.

Comparison of cardiovascular mortality and all-cause mortality. In the all-cause mortality analysis, maximal METs was the most significant predictor in both age groups. Exercise capacity is a reflection of the integrity of the cardiopulmonary system and a marker of a physically active lifestyle. In the elderly, patients with lower maximal METs achieved often have other comorbid conditions that limit their exercise capacity and likely lead to increased mortality. In the all-cause mortality data, resting ST depression decreases in significance. Thus, the findings reveal that the end points determine which variables will be most significantly predictive. Previous studies that used all-cause mortality as an end point have found METs to be the most significant predictor (9,34). When using an end point of cardiovascular mortality, resting ST depression and delta PRP are more prognostic. Lauer et al. (35) have pointed out the inaccuracy of death certificates in regards 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.

Comparison with prior studies. The largest previous study among the elderly came from the Rochester Epidemiology Project (9). They used the records of the doctor's offices performing treadmill testing to identify a retrospective cohort of Olmsted County residents tested between 1987 and 1989. During the study period, initial treadmill tests were performed in 2,593 persons <65 years of age and 514 persons ≥65 years of age. Peak METs were the only exercise-testing variable associated with all-cause mortality in both age groups, and the strength of association was similar. Metabolic equivalents and angina with exercise testing were associated with cardiac events in both age

Table 5. AUC of ROC Plots for ST Depression, DTS, and VA/UWV Score

	AUC ± SE		Z-Score for Age Group Comparison
	Above 65	Below 65	
ST depression	0.68 ± 0.019	0.66 ± 0.029	Nonsignificant
DTS	0.72 ± 0.018	0.72 ± 0.026	Nonsignificant
VA/WV score	0.79 ± 0.016	0.75 ± 0.025	Nonsignificant
Z-score ST depression vs. DTS	Nonsignificant	Nonsignificant	—
Z-score ST depression vs. VA/UWV score	Significant	Significant	—

AUC = area under the curve; DTS = Duke Treadmill Score; Nonsignificant = ROC curves are not significantly different (Z-score <1.96); ROC = receiver operating characteristic; Significant = ROC curves are significantly different (Z-score >1.96); VA/UWV score = Veterans Affairs/University of West Virginia score.

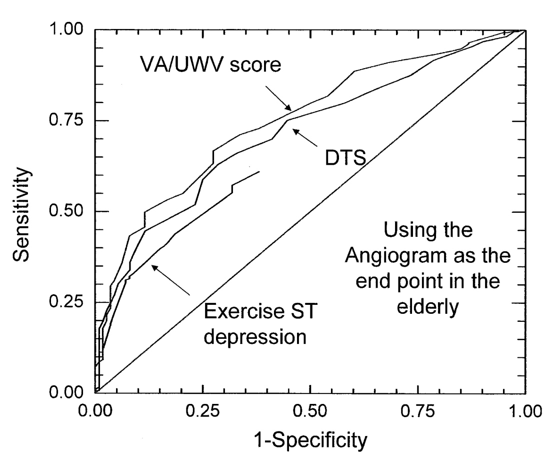

Figure 5. Receiver operator characteristic plot for the Veterans Affairs/ University of West Virginia (VA/UWV) and Duke Treadmill Scores (DTS) in patients ≥65 years of age with the end point being angiographic disease.

groups, whereas exercise-induced ST depression was associated with cardiac events only in younger persons. With consideration of clinical variables, peak METs was the only additional exercise variable that was predictive of death ($p < 0.001$) and cardiac events ($p < 0.05$); the strength of the association was similar in both age groups. Our study used cardiovascular death as the primary outcome to focus on the exercise treadmill test and included angiographic data.

Comparison of diagnostic characteristics of the scores/angiographic data. In the elderly, compared with the DTS, only the AUC of the VA/UWV angiographic diagnostic score was significantly greater than that of the ST response alone. Thus, although the DTS has been shown to have diagnostic utility for angiographic coronary disease, it is likely to be limited in the elderly. This is the first study to evaluate the diagnostic characteristics of the DTS in the elderly.

Strengths. This study defines more clearly the diagnostic and prognostic utility of exercise testing in elderly men. Strengths include a long follow-up, the inclusion of all consecutive referrals of elderly patients to two similar academically aligned Veterans Administration medical centers spanning over 12 years, an angiographic subset, and the application of standardized methodology. This study is the largest to date that documents the safety of "symptom- and sign-" limited maximal exercise testing in the elderly.

Study limitations. Our survival analysis is limited because we did not remove patients from observation based on whether they had a cardiovascular procedure during follow-up, because that information was not available. While the DTS was generated using the end point of cardiovascular death like this study (27,36), patients with interventions such as bypass surgery or catheter procedures had their follow-up ended when interventions occurred in the Duke study. Censoring for revascularization would have improved our predictive ability and allowed follow-up to the time of the cardiovascular procedure. The use of the California death registry was a minor limitation. Only eight patients died from California when we compared the deaths to those reported using the social security death index, and we considered these eight deaths to be cardiovascular causes.

Recently, there has been important data showing the utility of exercise treadmill data in the recovery period after the completion of exercise. The use of heart rate recovery, ST depression during recovery, and ventricular ectopy have all been shown to have prognostic importance (37,38). The absence of this data is a limitation, as there may be more predictive variables in a treadmill score. Another issue is in regards to the angiographic data using 50% stenosis to define significant angiographic CAD. It is appreciated now that a reduction in coronary flow reserve is associated with 70% stenosis. However, in a meta-analysis of 150 angiographic correlative studies, the study concluded that there were no appreciable differences in test characteristics between 50% and 70% diameter obstruction as a criterion for defining disease (39).

Another limitation is the reduction of the utility of the ST segment response by inclusion of exercise tests with low heart rates, baseline ECGs with up to 1.0 mm ST depression, and the higher prevalence of heart-rate-limiting drugs in the elderly. Furthermore, we are limited by selection bias, which is especially pertinent in the elderly.

Summary. Our findings demonstrate the importance of exercise testing as an important prognostic and diagnostic modality in the elderly. The exercise test offers additional value over clinical variables alone both for diagnosis and prognosis. This study demonstrates that exercise test scores can be helpful for the diagnosis and prognosis of coronary disease in the elderly.

Reprint requests and correspondence: Dr. Steve Lai, VA Medical Center Building 1, Box 181G, 4150 Clement Street, San Francisco, California 94121. E-mail: slai1@itsa.ucsf.edu.

REFERENCES

1. DeStefano F, Merritt RK, Anda RF, et al. Trends in nonfatal coronary heart disease in the United States, 1980 through 1989. Arch Intern Med 1993;153:2489–94.
2. Rodeheffer RJ, Gerstenblith G, Becker LC, et al. Exercise cardiac output is maintained with advancing age in healthy human subjects: cardiac dilatation and increased stroke volume compensate for a diminished heart rate. Circulation 1984;69:203–13.
3. Kasser IS, Bruce RA. Comparative effects of aging and coronary heart disease on submaximal and maximal exercise. Circulation 1969;39: 759–74.
4. Correia L, Lakatta E, O'Connor F, et al. Attenuated cardiovascular reserve during prolonged submaximal cycle exercise in healthy older subjects. J Am Coll Cardiol 2002;40:1290–7.
5. Kwok JM, Miller TD, Hodge DO, et al. Prognostic value of the Duke treadmill score in the elderly. J Am Coll Cardiol 2002;39:1475–81.
6. Fletcher GF, Froelicher VF, Hartley LH, et al. Exercise standards: a statement for health professionals from the American Heart Association. Circulation 1990;82:2286–321.
7. Gibbons RJ, Balady GJ, Beasley JW, et al. ACC/AHA guidelines for exercise testing. J Am Coll Cardiol 1997;30:260–311.
8. Spin J, Prakash M, Froelicher V, et al. The prognostic value of exercise testing in elderly men. Am J Med 2002;112:453–9.
9. Goraya T, Jacobsen S, Pellikka P, et al. Prognostic value of treadmill exercise testing in elderly persons. Ann Intern Med 2000;132:862–70.
10. Ustin J, Umann T, Froelicher V. Data management: a better approach. Physicians Comput 1994;12:30–3.
11. Froelicher V, Shiu P. Exercise test interpretation system. Physicians Comput 1996;14:40–4.
12. Wolthuis R, Froelicher VF, Fischer J, et al. New practical treadmill protocol for clinical use. Am J Cardiol 1977;39:697–700.
13. Myers J, Buchanan N, Walsh D, et al. A comparison of the ramp versus standard exercise protocols. J Am Coll Cardiol 1991;17:1334–42.
14. 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;73:591–6.
15. 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;112:11–6.
16. Reid M, Lachs M, Feinstein A. Use of methodological standards in diagnostic test research. JAMA 1995;274:645–51.
17. Yang JC, Wesley RC, Froelicher VF. Ventricular tachycardia during routine treadmill testing: risk and prognosis. Arch Intern Med 1991;151:349–53.
18. Raxwal V, Shetler K, Morise A, et al. A simple treadmill score to diagnose coronary disease. Chest 2001;119:1933–40.
19. Yusuf S, Zucker D, Peduzzi P, et al. Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomized

trials by the coronary artery bypass graft surgery trialists collaboration. Lancet 1994;344:563–70.
20. Lipinski M, Do D, Morise A, Froelicher V. What percent luminal stenosis should be used to define angiographic coronary artery disease for noninvasive test evaluation? Ann Noninvasive Electrocardiol 2002; 143:98–105.
21. 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;102:126–40.
22. 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;102:1193–209.
23. 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;282:1047–53.
24. Fearon WF, Lee DP, Froelicher VF. The effect of resting ST segment depression on the diagnostic characteristics of the exercise treadmill test. J Am Coll Cardiol 2000;35:1206–11.
25. Shaw LJ, Peterson ED, Shaw LK, et al. Use of a prognostic treadmill score in identifying diagnostic coronary disease subgroups. Circulation 1998;98:1622–30.
26. Fearon WF, Gauri AJ, Myers J, et al. A comparison of treadmill scores to diagnose coronary artery disease. Clin Cardiol 2002;25:117–22.
27. 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.
28. Gobel FL, Nordstrom LA, Nelson RR, et al. The rate-pressure product as an index of myocardial oxygen consumption during exercise in patients with angina pectoris. Circulation 1978;57:549–56.
29. Kitamura K, Jorgensen CR, Gobel FL, et al. Hemodynamic correlates of myocardial oxygen consumption during upright exercise. J Appl Physiol 1972;32:516–22.
30. Daida H, Allison TG, Squires YR, et al. Peak exercise blood pressure stratified by age and gender in apparently healthy subjects. Mayo Clin Proc 1996;71:445–52.
31. Saunamäki KI, Andersen JD. Post-myocardial infarction exercise testing: clinical significance of a left ventricular function index and ventricular arrhythmias. Acta Med Scand 1985;218:271–8.
32. Villella M, Villella A, Barlera S, et al. Prognostic significance of double product and inadequate double product response to maximal symptom-limited exercise stress testing after myocardial infarction in 6296 patients treated with thrombolytic agents. Am Heart J 1999;137:443–52.
33. Hammermeister KE, DeRouen TA, Dodge HT. Variables predictive of survival in patients with coronary disease: selection by univariate and multivariate analyses from the clinical, electrographic, exercise, arteriograhic, and quantitative angiographic evaluation. Circulation 1979; 59:421–30.
34. Myers J, Prakash M, Froelicher V, et al. Exercise capacity and mortality among men referred for exercise testing. N Engl J Med 2002;346:793–801.
35. Lauer MS, Blackstone EH, Youn JB, et al. Cause of death in clinical research. J Am Coll Cardiol 1999;34:618–20.
36. Mark DB, Shaw L, Harrell FE, et al. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. N Engl J Med 1991;325:849–53.
37. Cole CR, Blackstone EH, Pashkow FJ, et al. Heart-rate recovery immediately after exercise as a predictor of mortality. N Engl J Med 1999;341:1351–7.
38. Frolkis JP, Pother CE, Blackstone EH, et al. Frequent ventricular ectopy after exercise as a predictor of death. N Engl J Med 2003;348: 781–90.
39. Gianrossi R, Detrano R, Mulvihill D, et al. Exercise-induced ST depression in the diagnosis of coronary artery disease: a meta-analysis. Circulation 1989;80:87–98.