ORIGINAL ARTICLE

# Maximal Exercise Testing Variables and 10-Year Survival: Fitness Risk Score Derivation From the FIT Project

Haitham M. Ahmed, MD, MPH; Mouaz H. Al-Mallah, MD, MSc; John W. McEvoy, MB, BCh, BAO; Khurram Nasir, MD, MPH; Roger S. Blumenthal, MD; Steven R. Jones, MD; Clinton A. Brawner, PhD; Steven J. Keteyian, PhD; and Michael J. Blaha, MD, MPH

Abstract

Objective: To determine which routinely collected exercise test variables most strongly correlate with survival and to derive a fitness risk score that can be used to predict 10-year survival.

Patients and Methods: This was a retrospective cohort study of 58,020 adults aged 18 to 96 years who were free of established heart disease and were referred for an exercise stress test from January 1, 1991, through May 31, 2009. Demographic, clinical, exercise, and mortality data were collected on all patients as part of the Henry Ford ExercIse Testing (FIT) Project. Cox proportional hazards models were used to identify exercise test variables most predictive of survival. A "FIT Treadmill Score" was then derived from the b coefficients of the model with the highest survival discrimination.

Results: The median age of the 58,020 participants was 53 years (interquartile range, 45-62 years), and 28,201 (49%) were female. Over a median of 10 years (interquartile range, 8-14 years), 6456 patients (11%) died. After age and sex, peak metabolic equivalents of task and percentage of maximum predicted heart rate achieved were most highly predictive of survival (P<.001). Subsequent addition of baseline blood pressure and heart rate, change in vital signs, double product, and risk factor data did not further improve survival discrimination. The FIT Treadmill Score, calculated as [percentage of maximum predicted heart rate þ 12(metabolic equivalents of task) e 4(age) þ 43 if female], ranged from 200 to 200 across the cohort, was near normally distributed, and was found to be highly predictive of 10-year survival (Harrell C statistic, 0.811).

Conclusion: The FIT Treadmill Score is easily attainable from any standard exercise test and translates basic treadmill performance measures into a fitness-related mortality risk score. The FIT Treadmill Score should be validated in external populations.

 n Mayo Clin Proc. 2015;90(3):346-355

Cardiorespiratory fitness has been associated with a lower risk of cardiovascular and all-cause mortality in men and women.[1-5] This relationship is due, in part, to the dependent of body mass index as well as numerous noncardiovascular conditions including cancer, degenerative muscle conditions, frailty, and

From the Ciccarone Center for the Prevention

Heart Disease, Johns Hopkins Hospital, Baltimore, MD (H.M.A., J.W.M., K.N., R.S.B., S.R.J., M.J.B.); Department of Cardiac Imaging, King Abdulaziz Cardiac Center, Riyadh, Kingdom of Saudi Arabia (M.H.A.-M.); Center for Prevention and Wellness Research, Baptist Health Medical Group, Miami Beach, FL (K.N.);

Affiliations continued at the end of this article.

inverse associations between exercise patterns and obesity, diabetes, and the metabolic syndrome. Regular exercise also reduces systemic inflammation, triglyceride and apolipoprotein B levels, and thrombotic risk and increases high-density lipoprotein cholesterol levels.[6] Furthermore, regular exercise favorably impacts long-term heart rate (HR), blood pressure (BP), sympathetic nervous system activity, and myocardial and endothelial function.[7-9] Finally, there is mounting evidence that fitness is psychological well-being.[10,11]

Despite these benefits, much remains to be characterized with respect to which variables from a maximal exercise test best predict downstream cardiovascular disease (CVD) and mortality. Resting HR, maximum oxygen uptake, metabolic equivalents of task (METs), HR recovery, and peak BP are some of the variables that have been correlated with future CVD events or mortality.[4,12,13] The Duke Treadmill Score (DTS) and other scores include many of these variables and are designed primarily for predicting obstructive coronary artery disease (CAD) over the short term in high-risk patients with exercise-induced chest discomfort.[14,15] However, no validated tools are available for incorporating exercise test variables from otherwise unremarkable tests into a risk score aimed at estimating long-term survival.

Mayo Clin Proc. n March 2015;90(3):346-355 n http://dx.doi.org/10.1016/j.mayocp.2014.12.013



The purpose of this study was to determine which routinely collected maximal exercise test variables most strongly correlate with survival and to derive a fitness risk score that can be used to predict 10-year survival.

## PATIENTS AND METHODS

### Study Design

This study is based on data from the Henry Ford ExercIse Testing (FIT) Project, a retrospective cohort study aimed at investigating the implications ofcardiorespiratoryfitness oncardiovascular outcomes and mortality. The FIT Project is a registry of 69,885 consecutive patients who underwent physician-referred exercise treadmill stress testing at the Henry Ford Health System (HFHS) in metropolitan Detroit, Michigan, from January 1, 1991, through May 31, 2009. Because the methodological details of the FIT Project have been reported previously,[16] only a brief overview is provided in this article.

Exercisetest,medicalhistory,andmedication data were collected by exercise physiologists and nurses at the time of testing. Supporting clinical data were derived from the electronic medical record and administrative databases shared across HFHS. All patients were deidentified before analysis. The FIT Project was approved by the institutional review boards of HFHS and Johns Hopkins Hospital. Because data were deidentified,theinstitutionalreviewboardsdeemed this study exempt from individual consent.

### Study Cohort

Data were collected on all 69,885 adult patients aged18to96years who were referredbya physician for an exercise stress test for any indication from January 1, 1991, through May 31, 2009. Patients with known CAD before testing (n¼10,190) or known congestive heart failure (n¼877) were excluded. Patients with maximum predicted HR (MPHR) of 110% or higher were thought to have had tachyarrhythmias that may have distorted HR data so were also excluded (n¼798, 1% of the sample). After exclusions, thefinalstudycohort comprised58,020patients.

### Treadmill Testing

All patients were tested using the standard Bruce protocol.[17] For individuals who underwent multiple exercise tests during the study period, only the results from the first test were considered in the registry. In accordance with the American College of Cardiology/American Heart Association guidelines,[18] treadmill testing was terminated at the discretion of the supervising clinician for reasons that included serious arrhythmias, abnormal hemodynamic responses, ST-segment changes diagnostic for ischemia, or exercise-limited symptoms such as chest discomfort or dyspnea.

Resting HR and BP were recorded before the exercise test with the patient supine. The MPHR was determined by the formula 220 age (years). In addition to continuous electrocardiographic monitoring, BP was measured every

3 minutes during the test. The highest recorded HR and BP were considered the peak HR and BP. Exercise capacity, expressed in estimated METs, was calculated by the Quinton Instrument Company treadmill controller based on achieved speed and elevation.

### Follow-up and Event Adjudication

Patients were followed up for the occurrence of all-cause mortality. Mortality ascertainment was conducted in April 2013 using an algorithm for searching the Social Security Death Master File. A complete algorithmic search was completed in 57,732 of 58,020 patients (99.5%).

### Statistical Analyses

The Framingham risk score was calculated using the full 1998 equation by Wilson et al.[19] Cox proportional hazards models were developed to identify exercise test variables that correlated significantly with mortality ($P<.05$) after adjustment for age and sex. Likelihood ratio $c^2$ statistics were used to rank variables in order of added explanatory power. Hierarchical models were then constructed with improvement in model discrimination tested using the Harrell C statistic.[20] P values for the models were determined using the Somers D statistic.[21] Interaction between variables was tested. For variables with significant interaction, the C statistic was repeated with inclusion of the interaction term to the model in order to determine if this altered predictive power.

TheCoxproportionalhazardsmodelwiththe highest C statistic was used to derive a survival riskscore.Wescaledthebcoefficientstothenearest integer for use in the risk score. This methodology has been used successfully for multiple other landmark risk scores including the TIMI (Thrombolysis in Myocardial Infarction) and GRACE (Global Registry of Acute Coronary Events) scores for acute coronary syndrome, andtheCHADS$_2$ (chronicheartfailure,hypertension, advanced age, diabetes, and prior stroke/ transient ischemic attack) score for thromboembolisminatrialfibrillation ,amongothers.[21-24]

Bootstrapping was used for validation of our main equation. Bootstrapping is superior to data splitting into "derivation" and "validation" cohorts because it uses the entire sample for both derivation and validation, allowing better estimation of model performance. It is also not subject to the chance element of a single, possibly fortuitous, splitinthe data.Fromourstudysample, 200 bootstrap samples were drawn (sample size, 58,020), each at random and with replacement, yielding the same size as the original sample. Thus, the process generated samples that were representative of those that would be obtained if repeated sampling were carried out on the actual population.

Distribution of the FIT Treadmill Score was plotted in a histogram. Predicted 10-year survival as a function of the FIT Treadmill Score (modeled as a continuous variable) was derived from a logistic regression model and plotted using a local polynomial smoothed line plot. Survival based on FIT Treadmill Score cutoffs (score greater than 100, 0 to 100, 100 to 1, and less than 100) was then tabulated and plotted using Kaplan-Meier statistics. Overall discrimination of

the risk score was evaluated using the Harrell C statistic.

Because antihypertensive medications such as b-blockers can impact MPHR, further sensitivity analyses were performed. Harrell C statistics were calculated for patients with and without hypertension as well as for those taking and not taking b-blockers in order to compare survival discrimination of the FIT Treadmill Score in these subgroups.

The a level was <.05 (2-sided). All statistical analyses were performed with Stata statistical software, version 13.1 (StataCorp).

## RESULTS

### Descriptive Characteristics

Descriptive characteristics of the study cohort are presented in Table 1. After exclusions, the cohort included a total of 58,020 patients; 28,201 (49%) were female, 37,131 (64%) were white, 16,744 (29%) were black, and 4145 (7%) were other races. The indications for stress testing were chest pain in 29,667 (51%), shortness of breath in 5127 (9%), palpitations in 2074 (4%), dizziness in 1135 (2%), and other reasons in 20,017 (34%).

Over a median of 10 years (interquartile range [IQR], 8-14 years) and 635,771 personyears of follow-up, 6456 of the 58,020 patients (11%) died. The median peak METs was 10 (IQR, 7-10), and the median percentage of MPHR (%MPHR) achieved was 91% (IQR, 86%-96%). Per Cox proportional hazards regression, hazard ratio for death differed significantly (all P<.05) based on age, race, sex, exercise test variables, and cardiac risk factors (Table 1).

### Maximal Exercise Test Variables

Likelihood ratio $c^2$ and overall discrimination by C statistic for each hierarchical model of exercise test variables are presented in Table 2. After age and sex, METs and %MPHR were the variables most predictive of survival with a C statistic of 0.811 (Table 2, Model 4; P<.001). The addition of subsequent exercise test variables did not lead to a further significant increase to the C statistic. Addition of traditional CVD risk factors, family history, Framingham risk score, and medication use also did not add a significant increase to the C statistic beyond that for Model 4 (Supplemental Table, available online at http:// www.mayoclinicproceedings.org).

Sex and age were not found to have significant interaction (P¼.76), but sex and METs did (P<.001). An interaction term of gender*METs was included in Model 4 but did not impact the C statistic, which remained at 0.811. Finally, correlation between the 2 most significant exercise test variables of METs and %MPHR was compared using a correlation coefficients matrix, and both variables were found to be independent (correlation of 0.33). Therefore, per comparisons of predictive power using the Harrell C and Somers D statistics, and after ruling out significant colinearity and variable interaction, Model 4 was deemed to have the

TABLE 1. Descriptive Characteristics of the Study Cohort[a,b]

| Variable | Total cohort (N¼58,020) | Alive at study end (n¼51,564) | Dead at study end (n¼6456) | Hazard ratio for death (95% CI) | P value |
|---|---|---|---|---|---|
| Demographic characteristics | | | | | |
| Age (y) | 53 (45-62) | 52 (44-60) | 66 (56-74) | 1.09 (1.09-1.09) | <.001 |
| Race | | | | | |
| White | 37,131 (64) | 33,034 (64) | 4097 (63) | 5.41 (3.00-9.79) | .009 |
| Black | 16,744 (29) | 14,599 (28) | 2145 (33) | 6.41 (3.54-11.59) | <.001 |
| Asian (ref) | 607 (1) | 596 (1) | 11 (0.1) | 1.00 (1.00-1.00) | NA |
| Hispanic | 386 (1) | 351 (1) | 35 (0.5) | 4.54 (2.31-8.94) | <.001 |
| Native American | 329 (1) | 322 (1) | 7 (0.1) | 1.25 (0.48-3.21) | .65 |
| Other | 2823 (5) | 2662 (5) | 161 (2) | 3.31 (1.80-6.11) | <.001 |

| | | | | | |
|---|---|---|---|---|---|
| Sex | | | | | |
| Male (ref) | 29,819 (51) | 26,028 (50) | 3791 (59) | NA | NA |
| Female | 28,201 (49) | 25,536 (50) | 2665 (41) | 0.80 (0.76-0.84) | <.001 |
| BMI (kg/m$^2$) | 28.7 (25.3-32.8) | 28.7 (25.5-32.9) | 27.8 (24.8-31.4) | 0.98 (0.97-0.98) | <.001 |
| Diabetes | 10,483 (18) | 8723 (17) | 1760 (27) | 2.08 (1.97-2.20) | <.001 |
| Hypertension | 35,665 (61) | 30,646 (59) | 5019 (78) | 2.50 (2.36-2.65) | <.001 |
| Hyperlipidemia | 24,880 (43) | 22,459 (44) | 2421 (38) | 0.92 (0.87-0.96) | .001 |
| History of smoking | 24,061 (41) | 20,959 (41) | 3102 (48) | 1.35 (1.29-1.42) | <.001 |
| Family history of CHD | 29,842 (51) | 27,019 (52) | 2823 (44) | 0.72 (0.68-0.75) | <.001 |
| FRS | | | | | |
| (ref) <10% | 31,835 (55) | 30,005 (58) | 1830 (28) | 1.00 (1.00-1.00) | NA |
| 10%-20% | 15,087 (26) | 13,279 (26) | 1808 (28) | 2.05 (1.92-2.19) | <.001 |
| >20% | 8621 (15) | 6427 (12) | 2194 (34) | 4.69 (4.41-5.00) | <.001 |
| Cardiac medication use at baseline (n¼29,906)[c] | | | | | |
| Aspirin | 9997 (17) | 8346 (16) | 1651 (26) | 1.68 (1.59-1.78) | <.001 |
| Statin | 10,608 (18) | 9609 (19) | 999 (15) | 1.18 (1.10-1.27) | <.001 |
| Clopidogrel | 455 (1) | 375 (1) | 80 (1) | 2.11 (1.69-2.63) | <.001 |
| b-Blocker | 9776 (17) | 8310 (16) | 1466 (23) | 1.62 (1.53-1.72) | <.001 |
| ACEI/ARB | 10,903 (19) | 9453 (18) | 1450 (22) | 1.60 (1.51-1.70) | <.001 |
| Treadmill data | | | | | |
| Resting HR (bpm) | 72 (64-81) | 72 (64-81) | 72 (64-82) | 1.00 (1.00-1.00) | .09 |
| Resting SBP (mm Hg) | 130 (118-142) | 130 (114-140) | 140 (124-154) | 1.02 (1.02-1.02) | <.001 |
| Double product (mm Hg bpm) | 9348 (7952-10,920) | 9240 (7920-10,792) | 10,080 (8556-11,800) | 1.00 (1.00-1.00) | <.001 |
| Peak HR (bpm) | 153 (153-165) | 155 (142-166) | 135 (120-149) | 0.97 (0.97-0.97) | <.001 |
| Peak SBP (mm Hg) | 180 (160-198) | 180 (160-198) | 182 (162-202) | 1.00 (1.00-1.00) | .21 |
| METs achieved | 10 (7-10) | 10 (7-12) | 7 (5-10) | 0.75 (0.74-0.76) | <.001 |
| %MPHR achieved | 91 (86-97) | 92 (87-97) | 88 (79-95) | 0.96 (0.96-0.97) | <.001 |
| Achieved 85% MPHR | 46,489 (80) | 42,544 (82) | 3945 (61) | 0.39 (0.37-0.41) | <.001 |
| Achieved 90% MPHR | 33,483 (58) | 30,773 (60) | 2710 (42) | 0.47 (0.45-0.49) | <.001 |
| Achieved 100% MPHR | 7954 (14) | 7161 (14) | 793 (12) | 0.74 (0.69-0.80) | <.001 |
| Presenting symptom | | | | | |
| Chest pain | 29,667 (51) | 26,775 (52) | 2892 (45) | 1.67 (1.36-2.02) | <.001 |
| Shortness of breath | 5127 (9) | 4419 (9) | 708 (11) | 2.26 (1.85-2.77) | <.001 |
| Palpitations (ref) | 2074 (4) | 1965 (4) | 109 (2) | 1.00 (1.00-1.00) | NA |
| Dizziness/fatigue | 1135 (2) | 1011 (2) | 124 (2) | 2.24 (1.73-2.89) | <.001 |

[a]ACEI/ARB ¼ angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; BMI ¼ body mass index; bpm ¼ beats per minute; CHD ¼ coronary heart disease; FRS ¼ Framingham Risk Score; HR ¼ heart rate; METs ¼ metabolic equivalents of task; MPHR ¼ maximum predicted HR; NA ¼ not applicable; ref ¼ reference; SBP ¼ systolic blood pressure. [b]Data are presented as No. (percentage) of patients or median (interquartile range). [c]Complete medication data were not available on all patients.

highest survival discrimination and was used for risk score derivation.

### FIT Treadmill Score

The variables from Model 4 were included in a Cox proportional hazards model to yield the following Cox coefficients:

Survival ¼ 0:014ð%MPHR þ 0:182ðMETsÞ þ

0:6381ðfemale sexÞ 0:0613ðageÞ

Bootstrapping with 200 replications yielded the same coefficients with P<.001 for each of the coefficients and their SEs. The findings of identical coefficients from derivation as well as during bootstrap resampling indicated reliability and internal validity of the original

formula derived. The formula was then rescaled to integer coefficients by multiplying the equation by a constant (66.67) in order to achieve a FIT Treadmill Score with a range of numbers between 200 and 200 that is easy to remember and use clinically. The final FIT Treadmill Score was:

FIT Treadmill Score ¼ %MPHR þ 12ðMETsÞ

4ðageÞ þ 43 if female

A Cox proportional hazards model using the FIT Treadmill Score yielded a Harrell C statistic of 0.811, which was identical to the C statistic for Model 4.

The FIT Treadmill Score ranged from 200 to 200 for most of the patients (98.1%), with an IQR of 42 to 63. A histogram of the FIT Treadmill Score (Supplemental Figure 1, available online at http://www.mayoclinicproceedings.org) illustrates near normal distribution, with kurtosis of 2.81 and skew of 0.28. Figure 1 shows that estimated 10-year survival increased in a hyperbolic fashion as FIT Treadmill Score increased. The greatest change in survival slope was near score 0, with rapidly decreasing estimated survival for patients with negative scores.

Survival by FIT Treadmill Score is displayed by Kaplan-Meier curves in Figure 2. Age- and sex-stratified FIT Treadmill Scores are displayed by percentile in Table 3 along with 10-year survival by score. Patients with a score greater than 100 had 98% survival, those with a score of 0 to 100 had 97% survival, those with a score of 100 to 1 had 89% survival, and those with a score less than 100 had 62% survival at a median follow-up of 10 years.

TABLE 2. Hierarchical Models for Predictive Survival Based on Exercise Test Variables at Median Follow-up of 10 Years (in Decreasing Rank Order)

| Model | Model variables | LR $c^2$ | Harrell C statistic | P value |
|---|---|---|---|---|
| 1 | Age | 6591 | 0.766 | <.001 |
| 2 | Age þ sex | 6796 | 0.771 | <.001 |
| 3 | Age þ sex þ METs | 8471 | 0.807 | <.001 |
| 4 | Age þ sex þ METs þ %MPHR | 8656 | 0.811 | <.001 |
| | Model 4 þ change in HR | 8666 | 0.812 | <.001 |
| | Model 4 þ resting HR | 8661 | 0.812 | <.001 |
| | Model 4 þ 85% MPHR | 8674 | 0.811 | <.001 |
| | Model 4 þ 90% MPHR | 8700 | 0.811 | <.001 |
| | Model 4 þ 100% MPHR | 8656 | 0.811 | <.001 |
| | Model 4 þ indication for testing | 8658 | 0.811 | <.001 |
| | Model 4 þ double product | 8503 | 0.811 | <.001 |
| | Model 4 þ resting SBP | 8503 | 0.811 | <.001 |
| | Model 4 þ peak SBP | 8514 | 0.810 | <.001 |
| | Model 4 þ change in SBP | 8421 | 0.810 | <.001 |

HR ¼ heart rate; LR ¼ likelihood ratio; METs ¼ metabolic equivalents of task; %MPHR ¼ percentage of maximum predicted HR; SBP ¼ systolic blood pressure.

### Sensitivity Analyses

The Harrell C statistic for survival using the FIT Treadmill Score was 0.82 in patients without hypertension and 0.79 in patients with hypertension. The C statistic was 0.82 in patients not taking b-blockers and 0.76 in those taking b-blockers. The C statistic was 0.80 for blacks and 0.81 for whites.

## DISCUSSION

Our study is among the largest to date evaluating maximal exercise stress test variables and their relationship to survival. We observed several important findings. First, results from 58,020 adults referred for exercise stress testing revealed that after age and sex, METs achieved and %MPHR achieved were the exercise test variables most predictive of survival. Second, the FIT Treadmill Score, based on METs achieved and %MPHR, is easily attainable from any standard treadmill test, is near normally distributed, is quickly calculated, and has excellent discrimination for 10-year survival. Third, this score allows patients to compare their treadmill performance with that of their age- and sex-stratified peers and estimate differences in survival by performance. Fourth, once fitness was accounted for, traditional cardiovascular risk factors did not add incremental survival discrimination to our score (Supplement Table), which emphasizes the profound importance of cardiorespiratory fitness. Finally, our score

was documented to be reliable and internally valid by bootstrap resampling estimation.

Using Table 3, patients and physicians can (1) compare patients' FIT Treadmill Scores to age- and sex-stratified peers, (2) estimate likelihood of 10-year survival by score, and (3) estimate potential improvements in survival had treadmill performance been better. For example, a 60-year-old man with a score in the bottom 5th percentile of age- and sex-stratified scores can see that his estimated 10-year survival is 62% compared with 89% for his peers who were able to achieve median scores (Table 3, Supplemental Figure 2, A, available online at http://www.mayoclinicproceedings.org). Similarly, a 45-year-old woman with a score in the bottom 5th percentile of age- and sex-stratified scores can see that her estimated 10-year survival is 89% compared with 98% for her peers with scores in the top 10th percentile (Table 3, Supplement Figure 2, B). Thus, the FIT Treadmill Score could be a powerful tool for motivating behavioral change after stress testing.

In our cohort, peak METs was the exercise test variable most strongly predictive of survival. This finding is consistent with those of prior studies. Cardiorespiratory fitness has been found tobeassociatedwithsurvivalinmultiplelargecohorts including the Lipid Research Clinics studies, Cooper Center Longitudinal Study, Framingham Offspring Study, Copenhagen Male Study, and St James Women Take Heart Project, among others.[1-4,25-34] Indeed, this is the basis for the American College of Cardiology/American Heart Association emphasis on fitness categorization,inparticularinpatientswithpoorfunctional capacity (<4 METs) when identifying increased preoperative risk.[35]

After METs, the exercise test variable most predictive of survival was %MPHR achieved. This result is consistent with the findings of other investigators who have noted that % MPHR achieved correlates with mortality.[36,37] The normal HR response to exercise is based on the body's physiologic need to increase cardiac output. Conversely, chronotropic incompetence during exercise is an indication that the cardiac response to exercise is abnormal.[38]

We did not observe statistically significant incremental improvements to survival discrimination using resting systolic BP, peak systolic BP, or change (peak-rest) in systolic BP after adjusting for age, sex, METs, and % MPHR achieved (Table 2). We suspect that this lack of improvement may have been due, in part, to theeffects ofantihypertensivemedications that could decrease systolic BP without necessarily improving survival to the level of normotensive patients. Furthermore, resting BP was measured immediately before stress testing, so it is unclear how well this measurementreflectstrue restingBPthroughout the day.

We included age and sex in our fitness risk score because of both their independent, statistically significant correlations with mortality (P<.001) and the well-known interactions between age, sex, fitness, and cardiovascular outcomes. Other investigators have described discrepancies in chronotropic indices and

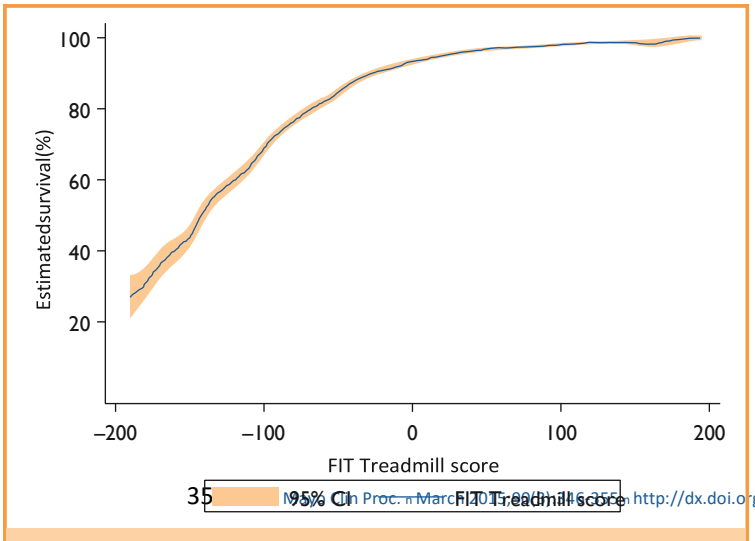


FIGURE 1. Estimated survival of 58,020 patients by FIT Treadmill Score over a median follow-up of 10 years.


Mayo Clin Proc. ■ March 2015;90(3):■■–■■ http://dx.doi.org/12.1016/j.mayocp.2014.12.013 www.mayoclinicproceedings.org

treadmill performance between men and women.[39] These discrepancies may be attributable to differences in muscle mass, physical activity characteristics, and sex hormones.[40-43] Moreover, differences in mortality based on sex have also been thoroughly documented. In the Western world, survival for women born between 2005 and 2010 is predicted to be 80.6 years compared with 75.6 years for men.[44] This difference is thought to be due to varying behavioral tendencies, biological traits, smoking prevalence, accident risk, and social characteristics, in addition to specifically.[45,46]

Similar to our FIT Treadmill Score, the DTS and other composite treadmill scoring systems have previously included METs.[14,15,47,48] In the original derivation of the DTS, logistic regression was used to predict severe stenosis (75%) or CAD with area under the receiver operating characteristic curve of 0.76.[14] Therefore, the DTS was designed to identify high-risk patients (ie, those with high pretest probability of coronary heart disease) in order to help guide the need for invasive angiography. However, the DTS has little value in lower-risk populations without abnormal test results. In fact, Gulati et al[49] reported that the DTS is no better at predicting long-term survival than exercise capacity alone in asymptomatic patients. Unlike the DTS, our FIT Treadmill Score is designed for lower-risk patients in whom the posttest probability of coronary heart disease is low. It emphasizes fitness and treadmill performance to predict long-term survival in lower-risk patients, rather than risk of obstructive CAD in high-risk patients.

Our study has several important strengths. First, with 635,771 person-years of follow-up, it is one of the largest studies to date evaluating exercise performance and subsequent mortality. Second, it included

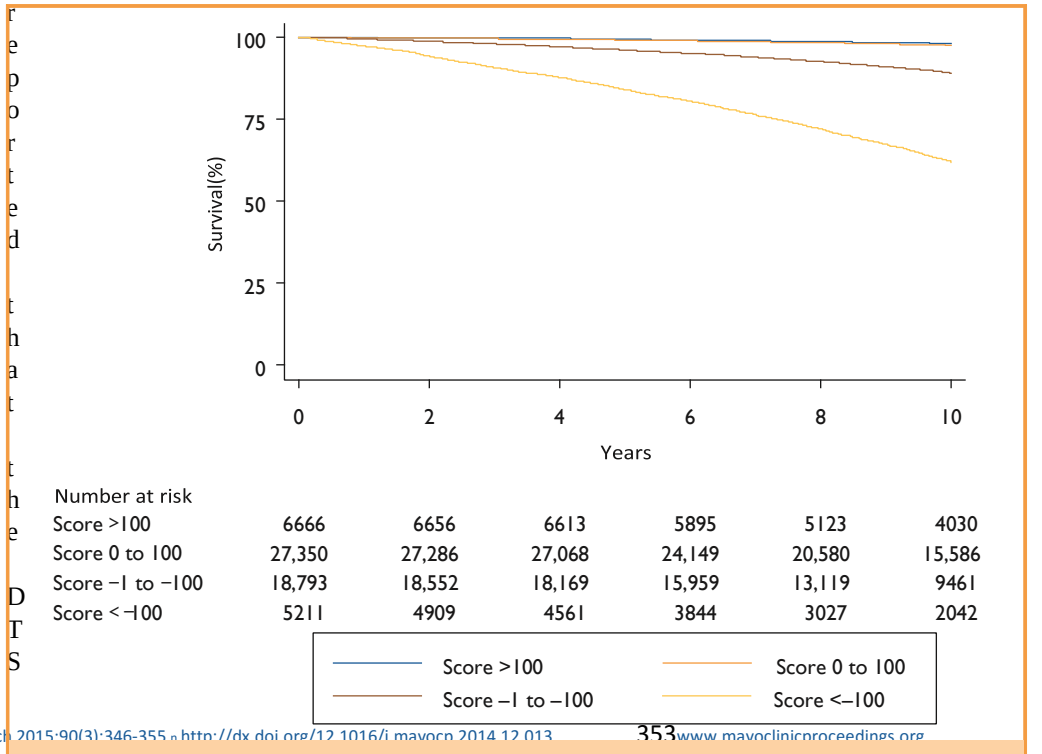


FIGURE 2. Kaplan-Meier survival estimates by FIT Treadmill Score (P<.001 by log rank test).

reported that the DTS discrepancies in cardiovascular mortality

both men and women of various race/ethnicity and age range beginning at 18 years, unlike most previous studies that have investigated only patients in older age groups. Thus, our results are more generalizable across various patient populations. Third, all patients were clinically referred for stress testing, so results regarding mortality prediction are highly relevant to patients referred for testing in routine clinical practice.

However, our study also some limitations. First, our results were from a single health system, and further studies will be needed to confirm our findings externally. Second, METs were estimated from the highest workload achieved as reported from the Quinton treadmill controller, which is how METs are calculated in many stress test laboratories including our own. It is unclear, however, if METs determination viathismethodisgeneralizabletoalternativeprotocols. Third, laboratory data were not available for all patients, which precluded adjustment for long-term markers of diabetes and hyperlipidemia. We repeated Model 4 in a subgroup of patients with available laboratory data (n¼10,735), including hemoglobin $A_{1c}$, total cholesterol, and high-density lipoprotein cholesterol values (Supplemental Table), and found that inclusion ofthesedatadidnotimprovesurvivaldiscrimination (C statistic, 0.791). Therefore, we do not believe that the absence of laboratory data has substantially harmed our score's predictive power, but comparison with a revised version that includes long-term markers of diabetes and hyperlipidemia remains necessary.

It is also important to note that b-blocker administration may negatively impact % MPHR and therefore may reduce the FIT
Treadmill Score's ability to predict survival in patients taking such medications. Our sensitivity analysis revealed that the Harrell C statistic decreased marginally from 0.82 in patients not taking b-blockers to 0.76 in patients taking b-blockers. Nonetheless, a C statistic of 0.76 still represents excellent survival discrimination comparable to that of the DTS, which also had an area under the receiver operating characteristic curve of 0.76.[14]

Finally, the FIT Project registry did not include comprehensive data regarding angina or ST-segment deviation. Such functional and electrocardiographic measures of ischemia are importantwhentryingtopredictCVDoutcomes, as is the goal of the DTS and other scores. It remains unclear, however, to what degree these measures would influence a risk score designed to use exercise test variables to predict longterm mortality. The DTS was designed for predicting obstructive CAD in high-risk patients, whereas our FIT Treadmill Score is designed to predict long-term survival in lower-risk patients

TABLE 3. Age- and Sex-Stratified FIT Treadmill Scores by Percentile in 58,020 Patients

| | FIT Treadmill Score | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Women | | | | | | | Men | | | | | | |
| Age | 5% | 10% | 25% | 50% | 75% | 90% | 95% | 5% | 10% | 25% | 50% | 75% | 90% | 95% |
| 30-34 | 65 | 78 | 99 | 122 | 140 | 164 | 170 | 40 | 68 | 81 | 112 | 125 | 144 | 149 |
| | 44 | 58 | 72 | | 114 | | 149 | | | | | 103 | 117 | 126 |
| 35-39 | 20 | | | 101 | | 143 | 120 | 14 | 36 | 59 | 90 | | | 105 |
| | | | | | | | 105 | | | | | | | |
| 40-44 | | 35 | 43 | 74 | 88 | 107 | | 7 | 10 | 38 | 66 | 82 | 94 | |
| 45-49 | 5 | 9 | 26 | 57 | 71 | 96 | | 31 | 17 | 15 | 31 | 60 | 71 | 82 |
| | | | | | | | | 55 | 41 | 10 | 7 | 38 | 50 | 58 |
| 50-54 | 27 | 14 | 5 | 33 | 50 | 67 | 84 | | | | | | | |
| | | | | | | | 52 | | | | | | | |
| 55-59 | 56 | 41 | 20 | 3 | | 263 | 21 | 84 | 69 | 48 | 18 | 7 | 26 | 32 |
| 60-64 | 84 | 68 | 46 28 | | | 8 1 1 5 | | 115 | 99 | 77 | 45 | 28 | 0 | 7 |
| 65-69 | 110 | 95 | | | 75 54 39 11 | | 3 | 148 129 105 | | | 86 | 55 | 37 17 | |
| | | | | | | | 28 | | | | | | | |
| 70-75 | 138 125 102 | | | 81 65 42 | | | | 172 158 136 115 | | | | 85 | 70 63 | |

| FIT Treadmill Score | >100 | 0 to 100 | 100 to 1 | Less than 100 |
|---|---|---|---|---|
| Survival at 10 years | 98% | 97% | 89% | 62% |

whose posttest probability of CHD is low. Regardless of these differences, the FIT Treadmill Score is a simple tool that can be used alongside the DTS and other scores to estimate survival, emphasize the profound importance of fitness, and hopefully help motivate behavioral change after exercise stress testing.

## CONCLUSION

We report a novel risk score for quantifying allcause mortality using exercise test variables. The FIT Treadmill Score [%MPHR þ 12(METs) e 4(age) þ 43 if female] is easily attainable from any standard treadmill test, is quickly calculated, has excellent discrimination for 10-year survival, and can be used to communicatefitness-relatedriskandmotivatebehavioral change after exercise stress testing. Our score was found to be reliable and internally valid in a large cohort of patients. External validation is needed at other institutions.

## SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at http://www.mayoclinicproceedings.org.

Abbreviations and Acronyms: BP = blood pressure; CAD = coronary artery disease; CVD = cardiovascular disease; DTS = Duke Treadmill Score; FIT = Henry Ford Exercise Testing; HFHS = Henry Ford Health System; HR = heart rate; IQR = interquartile range; METs = metabolic equivalents of task; MPHR = maximum predicted heart rate; %MPHR = percentage of MPHR

Affiliations (Continued from the first page of this article.):

Department of Medicine, Herbert Wertheim College of Medicine, Miami, FL (K.N.); Department of Epidemiology, Robert Stempel College of Public Health and Social Work, Miami, FL (K.N.); and Division of Cardiovascular Medicine, Henry Ford Hospital, Detroit, MI (M.H.A.-M., C.A.B., S.J.K.).

Correspondence: Address to Michael J. Blaha, MD, MPH, Ciccarone Center for the Prevention of Heart Disease, Johns Hopkins Hospital, Carnegie 565A, 600 N Wolfe St, Baltimore, MD 21287 (mblaha1@jhmi.edu).

## REFERENCES

1. Lee DC, Sui X, Artero EG, et al. Long-term effects of changes in cardiorespiratory fitness and body mass index on all-cause and cardiovascular disease mortality in men: the Aerobics Center Longitudinal Study. Circulation. 2011; 124(23):2483-2490.
2. Ahmed HM, Blaha MJ, Nasir K, et al. Low-risk lifestyle, coronary calcium, cardiovascular events, and mortality: results from MESA. Am J Epidemiol. 2013;178(1):12-21.
3. Berry JD, Willis B, Gupta S, et al. Lifetime risks for cardiovascular disease mortality by cardiorespiratory fitness levels measured at ages 45, 55, and 65 years in men: the Cooper Center Longitudinal Study. J Am Coll Cardiol. 2011;57(15): 1604-1610.
4. Mora S, Redberg RF, Cui Y, et al. Ability of exercise testing to predict cardiovascular and all-cause death in asymptomatic women: a 20-year follow-up of the Lipid Research Clinics Prevalence Study. JAMA. 2003;290(12):1600-1607.
5. Nes BM, Vatten LJ, Nauman J, Janszky I, Wisløff U. A simple nonexercise model of cardiorespiratory fitness predicts long-term mortality. Med Sci Sports Exerc. 2014;46(6): 1159-1165.
6. Ahmed HM, Blaha MJ, Nasir K, Rivera JJ, Blumenthal RS. Effects of physical activity on cardiovascular disease. Am J Cardiol. 2012; 109(2):288-295.
7. Whelton SP, Chin A, Xin X, He J. Effect of aerobic exercise on blood pressure: a meta-analysis of randomized, controlled trials. Ann Intern Med. 2002;136(7):493-503.
8. Duncan JJ, Farr JE, Upton SJ, Hagan RD, Oglesby ME, Blair SN. The effects of aerobic exercise on plasma catecholamines and blood pressure in patients with mild essential hypertension. JAMA. 1985;254(18):2609-2613.
9. Goto C, Higashi Y, Kimura M, et al. Effect of different intensities of exercise on endothelium-dependent vasodilation in humans: role of endothelium-dependent nitric oxide and oxidative stress. Circulation. 2003;108(5):530-535.
10. Ortega FB, Lee DC, Katzmarzyk PT, et al. The intriguing metabolically healthy but obese phenotype: cardiovascular prognosis and role of fitness. Eur Heart J. 2013;34(5):389-397.
11. Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, et al; American College of Sports Medicine. American College of Sports Medicine Position Stand: exercise and physical activity for older adults. Med Sci Sports Exerc. 2009;41(7):1510-1530.
12. Jensen MT, Suadicani P, Hein HO, Gyntelberg F. Elevated resting heart rate, physical fitness and all-cause mortality: a 16-year follow-up in the Copenhagen Male Study. Heart. 2013;99(12):882-887.
13. Skretteberg PT, Grundvold I, Kjeldsen SE, et al. Seven-year increase in exercise systolic blood pressure at moderate workload predicts long-term risk of coronary heart disease and mortality in healthy middle-aged men. Hypertension. 2013; 61(5):1134-1140.
14. Shaw LJ, Peterson ED, Shaw LK, et al. Use of a prognostic treadmill score in identifying diagnostic coronary disease subgroups. Circulation. 1998;98(16):1622-1630.
15. 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(12):849-853.
16. Al-Mallah MH, Keteyian SJ, Brawner CA, Whelton S, Blaha MJ. Rationale and design of the Henry Ford Exercise Testing Project (the FIT Project). Clin Cardiol. 2014;37(8):456-461.
17. Bruce RA, Pearson R, Lovejoy FW Jr, Yu PNG, Brothers GB. Variability of respiratory and circulatory performance during standardized exercise. J Clin Invest. 1949;28(6, pt 2):1431-1438.
18. Gibbons RJ, Balady GJ, Bricker JT, et al. ACC/AHA 2002 guideline update for exercise testing: summary article; a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1997 Exercise Testing Guidelines). J Am Coll Cardiol. 2002;40(8):1531-1540.
19. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97(18):1837-1847.
20. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982;247(18):2543-2546.
21. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somer's D. STATA J. 2010;10(3): 339-358.
22. Gage BF, van Walraven C, Pearce L, et al. Selecting patients with atrial fibrillation for anticoagulation: stroke risk stratification in patients taking aspirin. Circulation. 2004;110(16):2287-2292.
23. Eagle KA, Lim MJ, Dabbous OH, et al; GRACE Investigators. A validated prediction model for all forms of acute coronary syndrome: estimating the risk of 6-month postdischarge death in an international registry. JAMA. 2004;291(22):2727-2733.
24. Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI: a method for prognostication and therapeutic decision making. JAMA. 2000;284(7): 835-842.
25. Blair SN, Kohl HW III, Barlow CE, Paffenbarger RS Jr, Gibbons LW, Macera CA. Changes in physical fitness and allcause mortality: a prospective study of healthy and unhealthy men. JAMA. 1995;273(14):1093-1098.
26. Blair SN, Kohl HW III, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality: a prospective study of healthy men and women. JAMA. 1989;262(17):2395-2401.
27. Ekelund L-G, Haskell WL, Johnson JL, Whaley FS, Criqui MH, Sheps DS. Lipid Research Clinics Mortality Follow-up Study. Physical fitness as a predictor of cardiovascular mortality in asymptomatic North American men. N Engl J Med. 1988; 319(21):1379-1384.
28. Myers J, Prakash V, Froelicher D, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. N Engl J Med. 2002;346(11): 793-801.
29. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA. 2009;301(19):2024-2035.
30. Gulati M, Pandey DK, Arnsdorf MF, et al. Exercise capacity and the risk of death in women: the St James Women Take Heart Project. Circulation. 2003;108(13):1554-1559.
31. Stevens J, Cai J, Evenson KR, Thomas R. Fitness and fatness as predictors of mortality from all causes and from cardiovascular disease in men and women in the Lipid Research Clinics Study. Am J Epidemiol. 2002;156(9):832-841.
32. Hein HO, Suadicani P, Gyntelberg F. Physical fitness or physical activity as a predictor of ischemic heart disease? a

17-year follow-up in the Copenhagen Male Study. J Intern Med. 1992; 232(6):471-479.
33. Balady GJ, Larson MG, Vasan RS, Leip EP, O'Donnell CJ, Levy D. Usefulness of exercise testing in the prediction of coronary disease risk among asymptomatic persons as a function of the Framingham risk score. Circulation. 2004;110(14):19201925.
34. Wickramasinghe CD, Ayers CR, Das S, de Lemos JA, Willis BL, Berry JD. Prediction of 30-year risk for cardiovascular mortality
by fitness and risk factor levels: the Cooper Center Longitudinal Study. Circ Cardiovasc Qual Outcomes. 2014;7(4):597-602.
35. Fleisher LA, Beckman JA, Brown KA, et al. ACC/AHA guidelines on perioperative cardiovascular evaluation and care for noncardiac surgery: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 2002 Guidelines on Perioperative Cardiovascular Evaluation for Noncardiac Surgery): developed in collaboration with the American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Rhythm Society, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, and Society for Vascular Surgery [published corrections appear in Circulation. 2008;117(5):e154 and Circulation. 2008;118(9):e143-e144]. Circulation. 2007;116(17):e418-e499.
36. Leeper NJ, Dewey FE, Ashley EA, et al. Prognostic value of heart rate increase at onset of exercise testing. Circulation. 2007;115(4):468-474.
37. La Rovere MT, Bigger JT Jr, Marcus FI, Mortara A, Schwartz PJ; ATRAMI (Autonomic Tone and Reflexes After Myocardial
Infarction) Investigators. Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction. Lancet. 1998;351(9101):478-484.
38. Ellestad MH. Chronotropic incompetence: the implications of heart rate response to exercise (compensatory parasympathetic hyperactivity?). Circulation. 1996;93(8):1485-1487.
39. Gulati M, Shaw LJ, Thisted RA, Black HR, Bairey Merz CN, Arnsdorf MF. Heart rate response to exercise stress testing in asymptomatic women: the St James Women Take Heart Project. Circulation. 2010;122(2):130-137.
40. McTiernan A, Wu L, Chen C, et al; Women's Health Initiative Investigators. Relation of BMI and physical activity to sex hormones in postmenopausal women. Obesity (Silver Spring). 2006;14(9):1662-1677.
41. Bertone-Johnson ER, Tworoger SS, Hankinson SE. Recreational physical activity and steroid hormone levels in postmenopausal women. Am J Epidemiol. 2009;170(9):1095-1104.
42. Chan MF, Dowsett M, Folkerd E, et al. Usual physical activity and endogenous sex hormones in postmenopausal women: the European Prospective Investigation into CancereNorfolk population study. Cancer Epidemiol Biomarkers Prev. 2007; 16(5):900-905.
43. Aoyagi Y, Shephard RJ. Sex differences in relationships between habitual physical activity and health in the elderly: practical implications for epidemiologists based on pedometer/accelerometer data from the Nakanojo Study. Arch Gerontol Geriatr. 2013; 56(2):327-338.
44. United Nations, Department of Economic and Social Affairs. World Population Prospects: The 2012 Revision. United Nations website. http://esa.un.org/wpp/. Published 2012. Updated April 14, 2014. Accessed October 1, 2013.
45. Candore G, Balistreri RC, Colonna-Romano G, et al. Genderrelated immune-inflammatory factors, age-related diseases, and longevity. Rejuvenation Res. 2010;13(2-3):292-297.
46. Pan Z, Chang C. Gender and the regulation of longevity: implications for autoimmunity. Autoimmun Rev. 2012;11(6-7):A393A403.
47. Lai S, Kaykha A, Yamazaki T, et al. Treadmill scores in elderly men. J Am Coll Cardiol. 2004;43(4):606-615.
48. Prakash M, Myers J, Froelicher VF, et al. Clinical and exercise test predictors of all-cause mortality: results from >6,000 consecutive referred male patients. Chest. 2001;120(3):1003-1013.
49. Gulati M, Arnsdorf MF, Shaw LJ, et al. Prognostic value of the Duke Treadmill Score in asymptomatic women. Am J Cardiol.
2005;96(3):369-375.