



Published in final edited form as:

Circulation. 2009 June 30; 119(25): 3189–3197. doi:10.1161/CIRCULATIONAHA.109.848382.

Importance of Treadmill Exercise Time as an Initial Prognostic Screening Tool in Patients with Systolic Left Ventricular Dysfunction

Eileen Hsich, MD, Eiran Z. Gorodeski, MD, MPH, Randall C. Starling, MD, MPH, Eugene H. Blackstone, MD, Hemant Ishwaran, PhD, and Michael S. Lauer, MD

Department of Cardiovascular Medicine (EH, EZG, RCS), Department of Quantitative Health Sciences (HI) and Cardiothoracic Surgery (EHB) at the Cleveland Clinic and Case Western Reserve University School of Medicine (EH, RCS), Cleveland, OH and the Division of Prevention and Population Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland (MSL)

Abstract

Background—We sought to determine if treadmill exercise time may be of value as an initial prognostic screening tool in ambulatory patients with impaired systolic function referred for cardiopulmonary exercise testing.

Methods and Results—We studied 2,231 adult systolic heart failure patients (27% women) who underwent cardiopulmonary stress testing using a modified Naughton protocol. We assessed the value of treadmill exercise time for prediction of all-cause death and a composite of death or UNOS status 1 heart transplantation. During a mean follow up of 5 years, 742 (33%) patients died. There were 249 (11%) UNOS status 1 heart transplants. Treadmill exercise time was predictive of death and the composite outcome in both women and men, even after accounting for peak oxygen consumption and other clinical covariates (adjusted hazard ratio of lowest versus high sex-specific quartile for prediction of death 1.70, 95% CI 1.05–2.75, $P=0.03$, and for prediction of the composite outcome 1.75, 95% CI 1.15–2.66, $P=0.009$). For a one minute change in exercise time there was a 7% increased hazard of death (e.g. comparing 480 to 540 seconds HR 1.07, 95% CI 1.02–1.12, $P=0.004$).

Conclusions—Since cardiopulmonary stress testing is not available in every hospital, treadmill exercise time using a modified Naughton protocol may be of value as an initial prognostic screening tool.

Keywords

Heart failure; exercise; sex; prognosis

Correspondence: Please address all correspondence and requests for reprints to Michael S. Lauer MD, FACC, FAHA, Director, Division of Prevention and Population Science, National Heart, Lung, and Blood Institute, National Institutes of Health, Rockledge Center II, 6701 Rockledge Drive, Room 10122, Bethesda, MD 20892, Telephone#: 301-435-0422, Fax #: 301-480-1864, lauerm@nhlbi.nih.gov.

Disclosures: None

Peak oxygen consumption (VO_2) remains one of the most powerful single predictors of mortality for heart failure (HF) patients with severe systolic left ventricular dysfunction¹⁻³. Current guidelines suggest ambulatory patients be considered for transplantation when the peak VO_2 ≥ 14 ml/kg/min or ≥ 12 ml/kg/min in the setting of β -blockade³. Although cardiopulmonary exercise testing is routinely used to determine candidacy for heart transplantation, it is not known whether a simple measurement of treadmill exercise time may be of comparable value as an initial prognostic screening tool. We sought to evaluate whether (1) treadmill exercise time predicts survival in HF patients with systolic dysfunction, (2) treadmill exercise time adds incremental prognostic value beyond established risk factors including peak VO_2 and (3) a model with treadmill exercise time in addition to established risk factors better classifies risk than a model with established risk factors alone.

Methods

We performed an observational prospective cohort study at the Cleveland Clinic involving consecutive patients with left ventricular ejection fraction $< 40\%$ who underwent cardiopulmonary stress testing between August 1997 to April 2007. We focused specifically on patients who were tested according to the modified Naughton protocol, which was the most common protocol used in our laboratory for this type of patient. Patients were excluded if they were < 18 years old or had no United States Social Security number. The study was approved by the institutional review board at the Cleveland Clinic, and since all data were collected and recorded as part of routine clinical care the requirement for informed consent was waived.

Clinical Data

All demographic information, medications, medical and surgical history, heart rate, blood pressure, directly measured height and weight, and stress results were obtained at the time of stress testing and prospectively recorded in our electronic database. Biventricular pacemakers were not reported separately from other pacemakers during database entry. Glomerular Filtration Rate (GFR) was estimated using the Cockcroft-Gault equation⁴. Left ventricular ejection fraction was determined either by echocardiography, left ventriculography, or ECG-gated myocardial perfusion imaging. If a patient underwent more than one metabolic stress test using a modified Naughton protocol, only the first test was considered. Serum laboratory tests within 3 months of the cardiopulmonary study were included and only the laboratory tests closest to the stress test date were used.

Cardiopulmonary stress testing

Cardiopulmonary stress testing was symptom limited and patients were strongly advised to not use the handrails for support. Details regarding our laboratory's protocols have been previously published⁵. Briefly, the test was performed using a modified Naughton protocol which increases workload about 1 metabolic equivalent every 2 minutes⁶. Results were recorded on a MedGraphics cardiopulmonary system (St. Paul, Minnesota). Oxygen consumption, carbon dioxide production, minute ventilation, tidal volume, heart rate, blood pressure, and respiratory rate were obtained at rest, every 30 seconds during exercise, and

during recovery. Total duration of exercise, heretofore referred to as treadmill exercise time, was recorded *to the nearest second*.

End Points

The primary endpoints were all-cause mortality and a composite of all-cause mortality or United Network for Organ Sharing (UNOS) Status 1 heart transplantation during a mean follow-up of 5 years (maximum for survivors, 11 years). Mortality data was obtained by linking our database with the United States Social Security Administration Death Index. We previously validated this approach, yielding a sensitivity of 97%⁷, which is equivalent to a 97% follow-up rate. UNOS status 1 heart transplantation was based on Organ Procurement and Transplantation Network data as of June 18, 2008. All events (death or heart transplantation) were censored as of April 1, 2008.

Statistical Analysis

Baseline characteristics were reported as sex-specific quartiles of treadmill exercise time. Continuous variables were expressed as means with standard deviations and categorical variables were expressed as frequencies. Serum laboratory tests prior to October 1999 were systematically unavailable on our electronic database. We used informed imputation to fill in these missing values by constructing regression models for each laboratory measure on all other characteristics of our existing database except for outcome (10% of creatinine clearance, BUN, glucose, and sodium were imputed and 15% of serum hemoglobin values were imputed).

We generated Kaplan-Meier plots relating treadmill exercise time and cumulative mortality. We constructed unadjusted and adjusted Cox proportional hazards model to examine the association of treadmill exercise time and outcomes. Possible nonlinear associations were tested with restricted cubic splines. Potentially important interactions were tested. The proportional hazards assumption was tested by scaled Schoenfeld residuals and inspection of hazard ratio plots.

Model discrimination was assessed by calculating the out-of-bag c-index for time to event outcomes, an approach that we have used previously⁸. The out-of-bag (OOB) method involves obtaining bootstrap samples from the original cohort and using each sample to compute a prediction model. Each bootstrap sample left out about 37% of the data, which was referred to as the OOB data. The prediction model was applied to the OOB data to calculate the OOB c-index, a measure that is conceptually similar to the area under a receiver operating curve (AUC under ROC). The purpose of this method is to perform *multiple internal validations* with an internal cohort that has been randomly sampled multiple times. *This is a more conservative measure than what we have reported in previous work which was based on one assessment of the original data*⁹. Hence, we would expect that these OOB c-index values will be lower than what we previously reported, as they represent more conservative estimates.

To determine the change in prediction error attributable to each variable, we recalculated the prediction error after random permutation of that variable in the OOB data; a variable with a

high degree of importance would be expected to yield a greater change in the OOB c-index. The process was repeated 100 times for each variable.

To address clinical utility we constructed risk reclassification figures. We compared 1-year predicted risk estimates based on models of established risk factors with and without variables in question. We picked categories of risk of 15% based on the current 1-year mortality outcomes after cardiac transplantations¹⁰. We used two objective ways of quantifying improvement in categories as suggested by Pencina: net reclassification improvement (NRI) and integrated discrimination improvement (IDI)¹¹. The NRI and IDI were calculated by comparing individual patient's predicted risks at 1-year based on a Cox model 1 vs. another Cox model 2. These predicted risks were calculated using a modification of the R function `predict()` with censoring at 1-year.

Statistical analysis was performed using the SAS version 9.1.3 (SAS Institute Inc., Cary, NC) and R version 2.6.2 (www.R-project.org). We used Harrell's Design and Hmisc libraries for model construction and assessment, an R macro written by one of the authors (H.I.) for the OOB and change in prediction error analyses, and another R macro written by Katan and colleagues for calculating the observed risks in the reclassification figures. We used a SAS macro written by Pencina for calculating the NRI and IDI¹¹.

This work was supported in part by Health Resources and Services Administration contract 234-2005-370011C. The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. The authors had full access to the data and take responsibility for the integrity of the data. All authors have read the manuscript and approved it for submission.

Results

Our cohort consisted of 602 (27%) women and 1629 (73%) men. During a mean follow up of 5 years (maximum for survivors, 11 years), 155 women (26% of female cohort) and 587 (36% of male cohort) men died. There were 249 patients (16% women) who underwent heart transplantation as UNOS status 1. Only four patients were transplanted twice.

Table 1 shows the baseline characteristics of the 2231 patients according to sex-specific quartiles for treadmill exercise time. Compared to those in the highest quartile, those in the lowest quartile were older, more likely to have coronary artery disease, diabetes, implantable cardioverter-defibrillators, pacemakers, and had lower systolic blood pressures with higher resting heart rates. They also had lower peak VO_2 and lower peak respiratory exchange ratio. They were less likely to be on a β -blocker, or an angiotensin converting enzyme inhibitor and more likely to be on nitrates, hydralazine, angiotensin receptor blockers, diuretics and antiarrhythmic agents. Laboratory tests were similar except those in the lowest quartile had higher BUN and lower creatinine clearance.

Outcomes

Figure 1 shows the association of sex-specific quartiles of treadmill exercise time with all-cause mortality. The composite of death and UNOS status 1 heart transplant yielded similar sex-specific results (Appendix Figure 1). In both women and men treadmill exercise time predicted survival (death or death/UNOS status 1 transplantation) with the worst survival in the quartile with the lowest treadmill exercise time. After adjusting for all the variables listed in Table 1 including peak VO_2 in a Cox proportional hazards model, a shorter exercise time remained associated with an increase hazard of death (Table 2) as well as death and UNOS status 1 heart transplantation (Appendix Table 1). For a one minute change in exercise time there was a 7% increased hazard of death (e.g. for a comparison of 480 to 540 seconds HR 1.07, 95% CI 1.02–1.12, $P=0.004$). To determine if the imputed laboratory values affected our data, we performed a sensitivity analysis leaving out all laboratory variables and found similar results (HR 1.08, 95% CI 1.03–1.13, $P=0.002$).

We further evaluated the association of treadmill exercise time and survival for the subgroup of patients with peak $\text{VO}_2 \geq 14$ ml/kg/min. Even among these lower-risk patients, lower treadmill exercise time predicted worse outcomes in both women and men (Figure 2 and Appendix Figure 2). In this subset of patients with high peak VO_2 , a one minute change in exercise time also yielded a 7% increased hazard of death in a fully adjusted model (i.e., for a comparison of 480 to 540 seconds HR 1.07, 95% CI 1.00–1.14, $P=0.04$).

The c-index of the models containing all variables including exercise time were 0.723 for all-cause death and 0.744 for composite of death and UNOS status 1 heart transplantation. The out-of-bag (OOB) c-indexes, a more conservative measure, of the models containing all variables including treadmill exercise time were 0.698 (95% CI 0.670 to 0.726) for all-cause death (Table 3) and 0.727 (95% CI 0.703 to 0.751) for the composite outcome of death and UNOS status 1 heart transplantation (Appendix Table 2), indicating moderate discriminatory ability. There were no notable differences between the models that contained clinical variables and either peak VO_2 or treadmill exercise time or both. Treadmill exercise time was amongst the top three most important contributors to discriminative prediction in both sets of models (Figure 3 and Appendix Figure 3).

To better compare model performance within clinical categories, we classified patients into low-risk (<15%) and high-risk ($\geq 15\%$) categories of 1-year risk, with these cut-points based on the observed national 1-year survival after heart transplantation. We compared models with and without treadmill exercise time by cross-classifying predicted risks, stratified by whether or not the patients died (Figure 4), or developed the composite end-point of death or transplantation (Appendix Figure 4). If an additional variable were to add no predictive value, all points would fall on the line of identity. Spread around the line indicates modulation of predicted risk; if the variable correctly modulates predicted risk, there should be a greater preponderance of events (red open circles) above the line of identify. Clinical variables and usage of peak VO_2 (Figure 4A) or treadmill exercise time (Figure 4B) improved net reclassification (NRI) for the outcome of death and the composite outcome (Appendix Figure 4A–B). However, a particularly noteworthy finding is that there were no significant differences in net reclassification between a model with clinical variables and

treadmill exercise time versus clinical variables and peak VO_2 (Figure 4C and Appendix Figure 4C).

The integrated discrimination improvement (IDI) (a reclassification measure that is not limited to pre-determined categories of risk) for a model with treadmill exercise time was 1.9% ($P < 0.0001$) (Figure 4B) for the outcome of death and 3.4% ($P < 0.001$) (Appendix Figure 4B) for the composite outcome. There were no significant difference in IDI between a model with clinical variables and treadmill exercise time versus clinical variables and peak VO_2 (Figure 4C and Appendix Figure 4C).

Discussion

In a large cohort of patients with impaired left ventricular systolic function who underwent the same cardiopulmonary stress testing protocol, we found that treadmill exercise time predicts survival and yields similar prognostic value to peak oxygen consumption (peak VO_2). Even among low risk patients with a peak $\text{VO}_2 > 14$ ml/kg/min those who had a lower treadmill exercise time had markedly worse outcomes. Our findings are consistent with the hypothesis that exercise capacity is a valuable initial prognostic screening tool in patients with systolic left ventricular dysfunction.

Our study extends upon previous work demonstrating the prognostic power of exercise capacity^{12, 13}. Exercise capacity (i.e. treadmill time or “estimated” METs) is predictive of survival in both healthy adults and those with known coronary artery disease^{14–16}. Exercise tolerance reflects a number of prognostically important factors including cardiac function, endothelial function, pulmonary function, oxygen carrying capacity, and autonomic nervous system balance. However, its usefulness as a prognostic screening tool in patients with systolic left ventricular dysfunction is less established.

Treadmill exercise time has not been used in most heart failure survival models^{1, 17–19} but it has been previously explored in at least three heart failure studies. In the Captopril-Digoxin Research Group report, exercise time on a modified Naughton protocol was deemed to have no significant prognostic value. However, the cohort was relatively healthy as patients who exercised less than 240 seconds were excluded; the average exercise time was quite high at 558 seconds²⁰. In a small study of 60 heart failure patients there was a high degree of correlation between treadmill exercise time and peak VO_2 . However, due to the small sample size, this study did not have the power to determine prognostic significance²¹. More recently, a cohort of 31 children with idiopathic cardiomyopathy who underwent cardiopulmonary stress testing using a modified Naughton protocol were followed prospectively for a median follow-up of 1282 days and 20 of them died or underwent heart transplantation. Both univariable and multivariable analyses showed treadmill exercise time to be predictive of outcome²².

Can a simple exercise stress test replace cardiopulmonary stress testing for patients being evaluated for heart transplantation? Possibly but we do not advocate this approach. By both OOB change in c-index and IDI, we found that peak VO_2 is a more powerful predictor of mortality than treadmill exercise time. Other researchers have also found measured peak

VO₂ to be a better predictor than either estimated peak VO₂ or work rate^{23, 24}. Another advantage of cardiopulmonary stress testing includes the ability to measure patient effort (i.e. respiratory exchange ratio). When patients achieve an adequate level of effort, peak VO₂ is a reliable and reproducible measurement and does not depend on utilizing only one exercise protocol. Exercise time may not be as reproducible based on the results of Heart Failure and A Controlled Trial Investigation Outcomes of exercise training (HF-ACTION) which repeated within 7 days a baseline exercise study in 401 systolic heart failure patients (87% used a modified Naughton protocol) and noted that the peak VO₂ was unchanged but exercise time increased by 26 seconds on average²⁵. Cardiopulmonary stress testing can also help identify non-cardiac causes for shortness of breath. However, cardiopulmonary stress testing is not available in every hospital and therefore treadmill exercise time may be a reasonable screening tool for physicians to determine prognosis and need for referral to advanced heart failure treatment centers. Based on our data, less than 5 minutes and 17 seconds for females and less than 6 minutes for males may be useful treadmill exercise time cut-off values when using a modified Naughton protocol since in our cohort this corresponded with about a 15% one year mortality. Future research in other cohorts will be needed to verify our findings.

The strengths of our study include a large sample size with a large number of hard events, prospective data collection, adequate power to present sex-specific results^{26, 27}, and the recording of treadmill exercise time *to the nearest second*. All endpoints were determined by query of databases outside of our center. Deaths were determined by Social Security files, whereas UNOS 1 transplantation was determined by direct query of UNOS files. Of note, 20% of patients who experienced UNOS 1 transplantation were transplanted outside of our center. Thus, an ascertainment bias based on only considering events known to Cleveland Clinic information systems was avoided. Body mass index was calculated based on direct measures of height and weight. We adjusted for over 35 clinical/demographic variables which included sex, peak VO₂, and coronary artery disease and still found a longer exercise time interval to be associated with a lower hazard of death. We also used both traditional (C index) and contemporary statistical discrimination and reclassification methods¹¹ to assess model performance.

The C-index values reported in this manuscript are lower than what we reported earlier regarding a similar cohort⁹. We now have adopted a more sophisticated and conservative method (OOB c-index) than what we used previously. This approach effectively simulates numerous internal validations and as expected model performance would be worse if applied to the entire original dataset.

Limitations include that this is a single center study with only patients who could ambulate on a treadmill. Since the cohort includes patients from 1997, there are few patients using aldosterone antagonists, modest β -blockers usage, and about 25% with an ICD. However, these all compare favorably to what is seen in registries recently published²⁸⁻³¹.

Biventricular pacemakers were not reported separately during database entry but most are identified in the ICD category since at our institution biventricular pacemakers are almost always implanted with an ICD. Therefore, when adjusting for all possible confounding variables we are including those with an ICD alone and with an ICD combined with a

biventricular pacemaker. Laboratory tests like B-type natriuretic peptides were not included in our baseline characteristics since not routinely obtained at our center between 1997 and 2007. Although BNP has prognostic value it also has not been used in other heart failure survival models^{17, 32–35}. Like all epidemiologic studies we also cannot account for variables that change with time that may impact on death such as patients receiving an ICD after the cardiopulmonary stress test or changes in medication with time. However, this limitation is inherent in all studies that have confirmed the value of cardiopulmonary stress testing as a prognostic tool. We also did not compare treadmill exercise time to indices of ventilatory inefficiency such as the VE/VCO₂ slope, PetCO₂, oscillatory ventilation, and oxygen uptake efficiency slope which are more powerful risk predictors when compared with peak VO₂. These indices require minute-by-minute measurements which was not available in our database.

Despite these limitations, we found that treadmill exercise time is a powerful and independent predictor of outcomes in both women and men with impaired left ventricular systolic function, even after accounting for peak oxygen consumption. These findings have potentially important implications for our understanding of the clinical pathophysiology of systolic cardiac dysfunction as well as for finding optimal pathways to screen heart failure patients for possible transplantation. Since cardiopulmonary stress testing is not available in every hospital, treadmill exercise time using a modified Naughton protocol may be of value as an initial prognostic screening tool.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources: Supported by American Heart Association Scientist Development Grant 0730307N and National Heart, Lung, and Blood Institute CAN #8324207

References

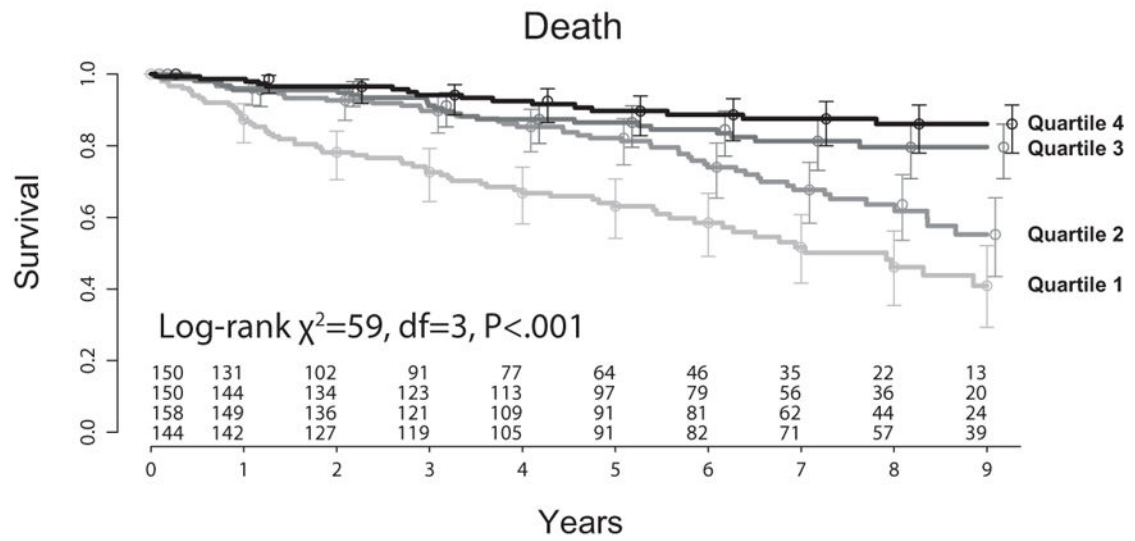
1. Mancini DM, Eisen H, Kussmaul W, Mull R, Edmunds LH Jr, Wilson JR. Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. *Circulation*. 1991; 83:778–786. [PubMed: 1999029]
2. Hunt SA, Abraham WT, Chin MH, Feldman AM, Francis GS, Ganiats TG, Jessup M, Konstam MA, Mancini DM, Michl K, Oates JA, Rahko PS, Silver MA, Stevenson LW, Yancy CW, Antman EM, Smith SC Jr, Adams CD, Anderson JL, Faxon DP, Fuster V, Halperin JL, Hiratzka LF, Jacobs AK, Nishimura R, Ornato JP, Page RL, Riegel B. ACC/AHA 2005 Guideline Update for the Diagnosis and Management of Chronic Heart Failure in the Adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. *Circulation*. 2005; 112:e154–235. [PubMed: 16160202]
3. Mehra MR, Kobashigawa J, Starling R, Russell S, Uber PA, Parameshwar J, Mohacsi P, Augustine S, Aaronson K, Barr M. Listing criteria for heart transplantation: International Society for Heart and Lung Transplantation guidelines for the care of cardiac transplant candidates—2006. *J Heart Lung Transplant*. 2006; 25:1024–1042. [PubMed: 16962464]

4. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron*. 1976; 16:31–41. [PubMed: 1244564]
5. Robbins M, Francis G, Pashkow FJ, Snader CE, Hoercher K, Young JB, Lauer MS. Ventilatory and heart rate responses to exercise : better predictors of heart failure mortality than peak oxygen consumption. *Circulation*. 1999; 100:2411–2417. [PubMed: 10595953]
6. Kleber FX, Waurick P, Winterhalter M. CPET in heart failure. *European Heart Journal Supplements*. 2004; 6:D1–D4.
7. Nishime EO, Cole CR, Blackstone EH, Pashkow FJ, Lauer MS. Heart rate recovery and treadmill exercise score as predictors of mortality in patients referred for exercise ECG. *Jama*. 2000; 284:1392–1398. [PubMed: 10989401]
8. Lauer MS, Martino D, Ishwaran H, Blackstone EH. Quantitative measures of electrocardiographic left ventricular mass, conduction, and repolarization, and long-term survival after coronary artery bypass grafting. *Circulation*. 2007; 116:888–893. [PubMed: 17679612]
9. O'Neill JO, Young JB, Pothier CE, Lauer MS. Peak oxygen consumption as a predictor of death in patients with heart failure receiving beta-blockers. *Circulation*. 2005; 111:2313–2318. [PubMed: 15867168]
10. UNOS Heart Kaplan-Meier graft survival rates for transplants performed 1997–2004. Jun 27. 2008 Available at: <http://www.optn.org/latestData/rptStrat.asp>. Accessed July 3, 2008
11. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008; 27:157–172. discussion 207–112. [PubMed: 17569110]
12. Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. *N Engl J Med*. 2002; 346:793–801. [PubMed: 11893790]
13. Gulati M, Pandey DK, Arnsdorf MF, Lauderdale DS, Thisted RA, Wicklund RH, Al-Hani AJ, Black HR. Exercise capacity and the risk of death in women: the St James Women Take Heart Project. *Circulation*. 2003; 108:1554–1559. [PubMed: 12975254]
14. Mora S, Redberg RF, Cui Y, Whiteman MK, Flaws JA, Sharrett AR, Blumenthal RS. 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; 290:1600–1607. [PubMed: 14506119]
15. Blair SN, Kampert JB, Kohl HW 3rd, Barlow CE, Macera CA, Paffenbarger RS Jr, Gibbons LW. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *Jama*. 1996; 276:205–210. [PubMed: 8667564]
16. Adabag AS, Grandits GA, Prineas RJ, Crow RS, Bloomfield HE, Neaton JD. Relation of heart rate parameters during exercise test to sudden death and all-cause mortality in asymptomatic men. *Am J Cardiol*. 2008; 101:1437–1443. [PubMed: 18471455]
17. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, Anand I, Maggioni A, Burton P, Sullivan MD, Pitt B, Poole-Wilson PA, Mann DL, Packer M. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation*. 2006; 113:1424–1433. [PubMed: 16534009]
18. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *Jama*. 2005; 293:572–580. [PubMed: 15687312]
19. Aaronson KD, Schwartz JS, Chen TM, Wong KL, Goin JE, Mancini DM. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation*. 1997; 95:2660–2667. [PubMed: 9193435]
20. Gradman A, Deedwania P, Cody R, Massie B, Packer M, Pitt B, Goldstein S. Predictors of total mortality and sudden death in mild to moderate heart failure. Captopril-Digoxin Study Group. *J Am Coll Cardiol*. 1989; 14:564–570. discussion 571–562. [PubMed: 2768707]
21. Clark AL, Coats AJ. Exercise endpoints in patients with chronic heart failure. *Int J Cardiol*. 2000; 73:61–66. [PubMed: 10748312]
22. Guimaraes GV, d'Avila VM, Camargo PR, Moreira LF, Luces JR, Bocchi EA. Prognostic value of cardiopulmonary exercise testing in children with heart failure secondary to idiopathic dilated

- cardiomyopathy in a non-beta-blocker therapy setting. *Eur J Heart Fail.* 2008; 10:560–565. [PubMed: 18486551]
23. Myers J, Gullestad L, Vagelos R, Do D, Bellin D, Ross H, Fowler MB. Clinical, hemodynamic, and cardiopulmonary exercise test determinants of survival in patients referred for evaluation of heart failure. *Ann Intern Med.* 1998; 129:286–293. [PubMed: 9729181]
 24. Myers J, Gullestad L, Vagelos R, Do D, Bellin D, Ross H, Fowler MB. Cardiopulmonary exercise testing and prognosis in severe heart failure: 14 mL/kg/min revisited. *Am Heart J.* 2000; 139:78–84. [PubMed: 10618566]
 25. Bensimhon DR, Leifer ES, Ellis SJ, Fleg JL, Keteyian SJ, Pina IL, Kitzman DW, McKelvie RS, Kraus WE, Forman DE, Kao AJ, Whellan DJ, O'Connor CM, Russell SD. Reproducibility of peak oxygen uptake and other cardiopulmonary exercise testing parameters in patients with heart failure (from the Heart Failure and A Controlled Trial Investigating Outcomes of exercise traiNing). *Am J Cardiol.* 2008; 102:712–717. [PubMed: 18773994]
 26. Hsich E, Chadalavada S, Krishnaswamy G, Starling RC, Pothier CE, Blackstone EH, Lauer MS. Long-term prognostic value of peak oxygen consumption in women versus men with heart failure and severely impaired left ventricular systolic function. *Am J Cardiol.* 2007; 100:291–295. [PubMed: 17631085]
 27. Elmariah S, Goldberg LR, Allen MT, Kao A. Effects of gender on peak oxygen consumption and the timing of cardiac transplantation. *J Am Coll Cardiol.* 2006; 47:2237–2242. [PubMed: 16750689]
 28. Yancy CW, Lopatin M, Stevenson LW, De Marco T, Fonarow GC. Clinical presentation, management, and in-hospital outcomes of patients admitted with acute decompensated heart failure with preserved systolic function: a report from the Acute Decompensated Heart Failure National Registry (ADHERE) Database. *J Am Coll Cardiol.* 2006; 47:76–84. [PubMed: 16386668]
 29. Gheorghiane M, Abraham WT, Albert NM, Greenberg BH, O'Connor CM, She L, Stough WG, Yancy CW, Young JB, Fonarow GC. Systolic blood pressure at admission, clinical characteristics, and outcomes in patients hospitalized with acute heart failure. *Jama.* 2006; 296:2217–2226. [PubMed: 17090768]
 30. Fonarow GC, Stough WG, Abraham WT, Albert NM, Gheorghiane M, Greenberg BH, O'Connor CM, Sun JL, Yancy CW, Young JB. Characteristics, treatments, and outcomes of patients with preserved systolic function hospitalized for heart failure: a report from the OPTIMIZE-HF Registry. *J Am Coll Cardiol.* 2007; 50:768–777. [PubMed: 17707182]
 31. Ezekowitz JA, Lee DS, Tu JV, Newman AM, McAlister FA. Comparison of one-year outcome (death and rehospitalization) in hospitalized heart failure patients with left ventricular ejection fraction >50% versus those with ejection fraction <50%. *Am J Cardiol.* 2008; 102:79–83. [PubMed: 18572040]
 32. Brophy JM, Dagenais GR, McSherry F, Williford W, Yusuf S. A multivariate model for predicting mortality in patients with heart failure and systolic dysfunction. *Am J Med.* 2004; 116:300–304. [PubMed: 14984814]
 33. Zugck C, Kruger C, Kell R, Korber S, Schellberg D, Kubler W, Haass M. Risk stratification in middle-aged patients with congestive heart failure: prospective comparison of the Heart Failure Survival Score (HFSS) and a simplified two-variable model. *Eur J Heart Fail.* 2001; 3:577–585. [PubMed: 11595606]
 34. Stempfle HU, Alt A, Stief J, Siebert U. The Munich score: a clinical index to predict survival in ambulatory patients with chronic heart failure in the era of new medical therapies. *J Heart Lung Transplant.* 2008; 27:222–228. [PubMed: 18267231]
 35. Lund LH, Aaronson KD, Mancini DM. Validation of peak exercise oxygen consumption and the Heart Failure Survival Score for serial risk stratification in advanced heart failure. *Am J Cardiol.* 2005; 95:734–741. [PubMed: 15757599]

Physicians often refer patients with advanced heart failure for exercise testing with metabolic gas exchange measurements in order to assess suitability for cardiac transplantation. Peak oxygen consumption and other metabolic measures are known to be powerful predictors of mortality. We asked whether treadmill exercise time according to a standardized protocol might be comparable to peak oxygen consumption for assessment of prognosis. We analyzed the outcomes of 2231 patients with systolic heart failure who all underwent metabolic exercise stress testing on a modified Naughton protocol. During a mean follow-up of 5 years, 742 patients (33%) died and 249 patients (11%) underwent heart transplantation urgently. We found that, after accounting for baseline clinical characteristics, treadmill exercise time performed similar to peak oxygen consumption for predicting poor outcome. For a one minute change in exercise time there was a 7% increased hazard of death (e.g. comparing 480 to 540 seconds HR 1.07, 95% CI 1.02–1.12, P=0.004. Mortality rates were particularly high for women who exercised less than 5 minutes and 17 seconds and for men who exercised less than 6 minutes. Our findings suggest that treadmill exercise time may be valuable as an initial prognostic screening tool in patients with advanced heart failure.

A. Women



B. Men

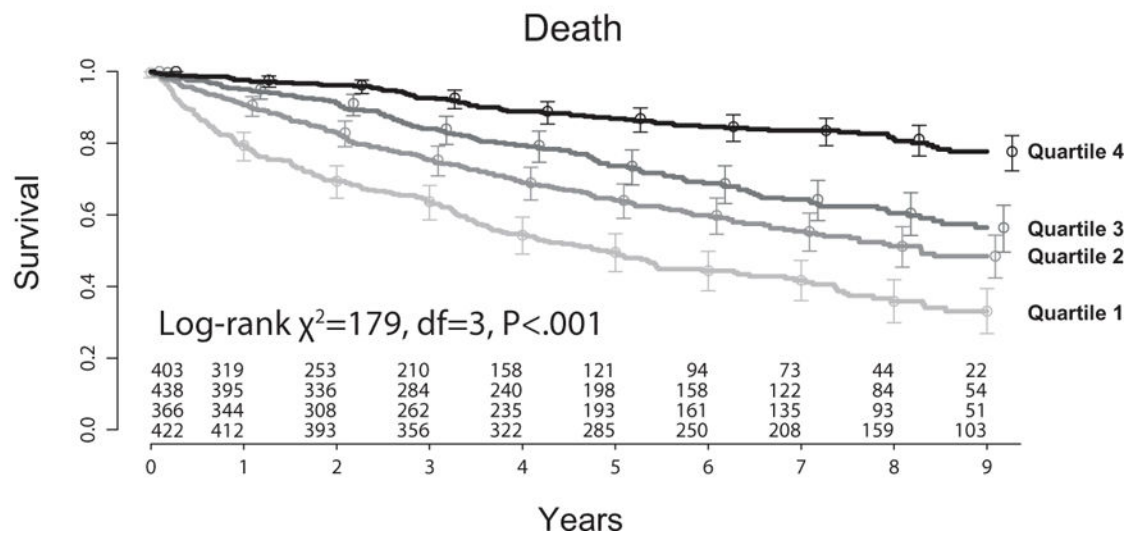
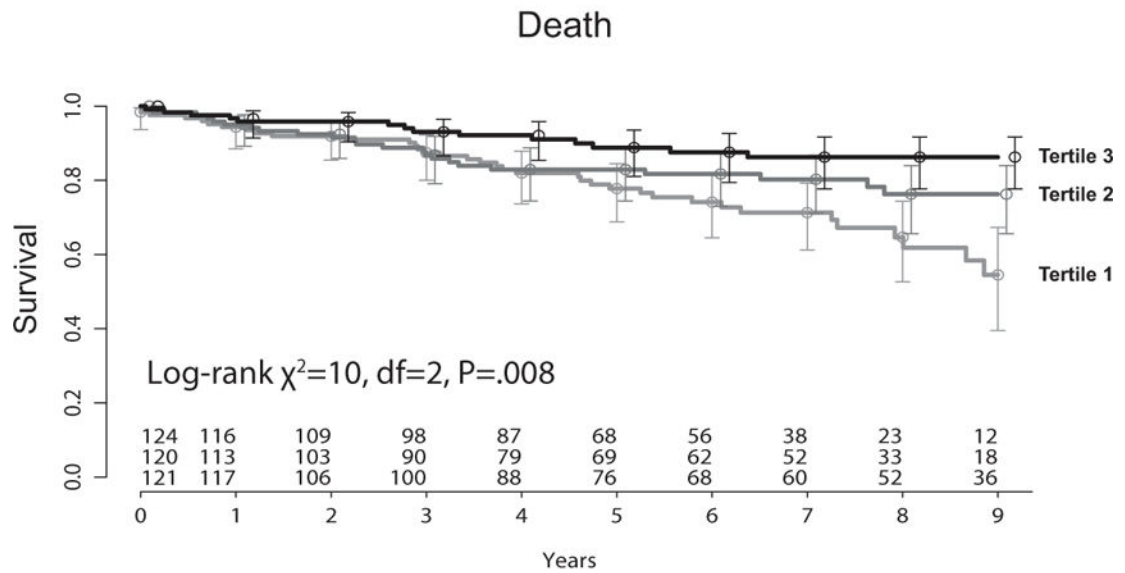


Figure 1. Kaplan-meier plots for (A) women and (B) men, stratified by quartiles of treadmill exercise time. (A) Women (N=602) Quartile 1: 21–315 seconds, Quartile 2: 317–478 seconds, Quartile 3: 480–600 seconds, Quartile 4: 603–1320 seconds. (B) Men (N=1629) Quartile 1: 35–357 seconds, Quartile 2: 360–480 seconds, Quartile 3: 484–657 seconds, Quartile 4: 660–1415 seconds.

A. Women, Peak VO₂ ≥14



B. Men, Peak VO₂ ≥14

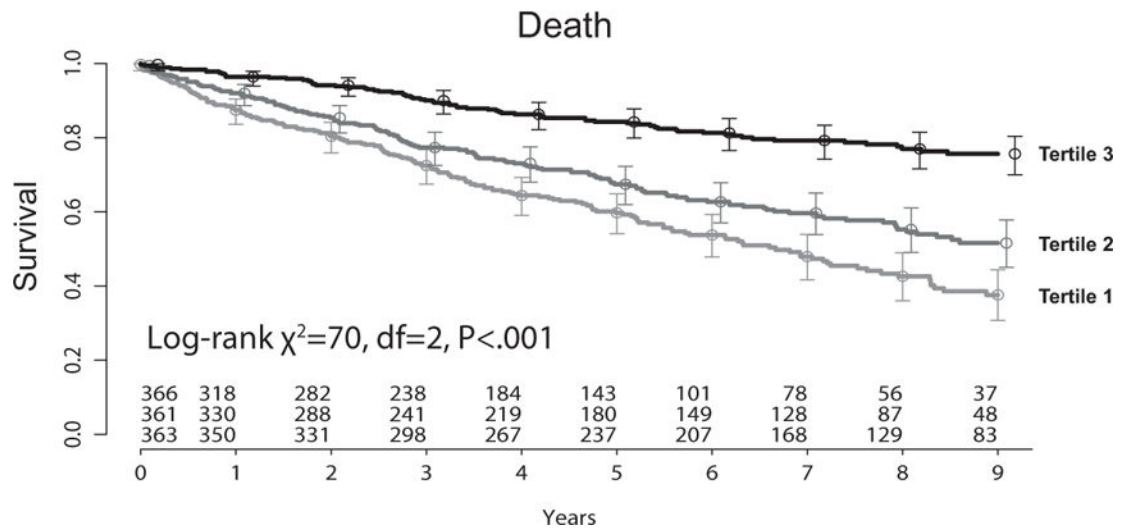


Figure 2. Kaplan-meier plots for (A) women and (B) men with peak VO₂ ≥ 14, stratified by tertiles of increasing treadmill exercise time. (A) Women: Tertile 1: < 518 seconds, Tertile 2: 518–639 seconds, Tertile 3: 640 seconds. (B) Men: Tertile 1: < 511 seconds, Tertile 2: 511–682 seconds, Tertile 3: 683 seconds. Peak VO₂= peak oxygen consumption

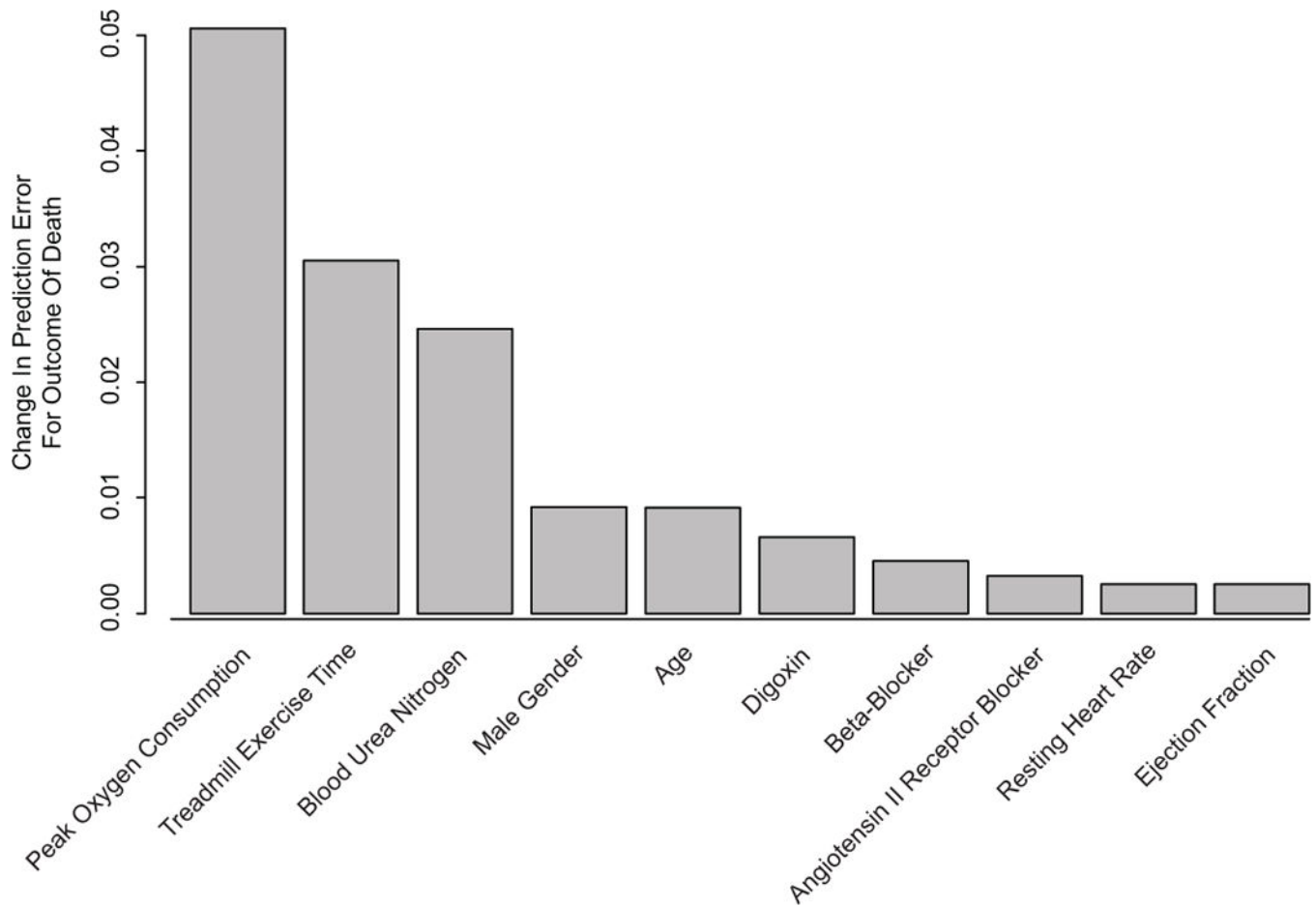
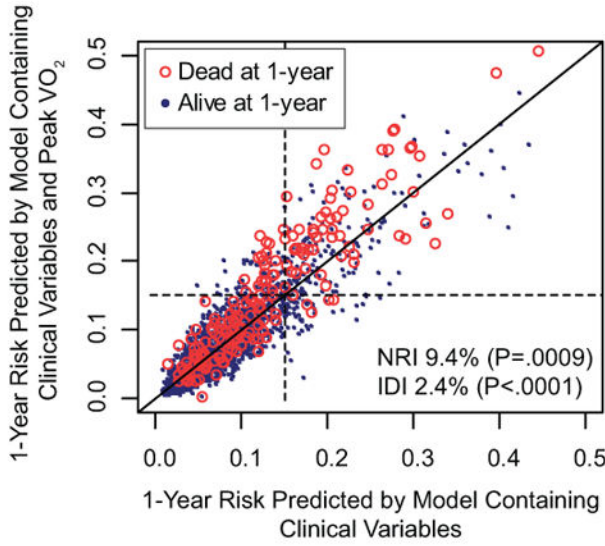
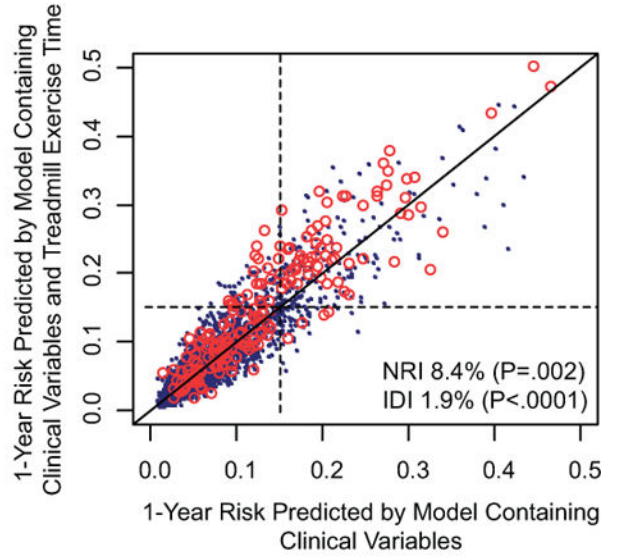


Figure 3. Change in out-of-bagging determined prediction error. Only the 10 most important variables are shown. Results are based on 100 bootstrapped samples.

A.



B.



C.

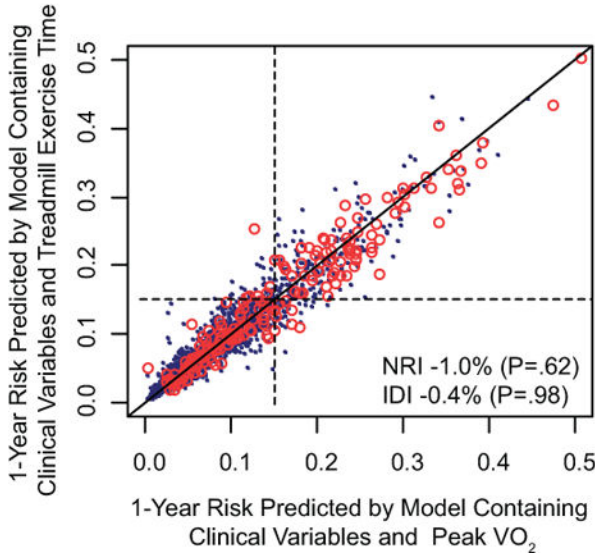


Figure 4.

Reclassification: One year low-risk (<15%) and high-risk (≥15%) survival categories are cut-offs based on the observed national 1-year survival after heart transplantation. The endpoint is death and improvement of classification is expressed as net reclassification improvement (NRI) which is dependent on the pre-defined cut-offs and integrated discrimination improvement (IDI) which is not limited to pre-determined risk categories. If a variable adds no predictive value to the model, all points fall on the dark line of identity within each figure. Spread around the line indicates modulation of predicted risk; if the variable correctly modulates predicted risk, there should be a greater preponderance of events (red open circles) above the line of identity. A. Model with clinical variables and

peak VO₂ compared to model with only clinical variables. Peak VO₂ with cardiac risk factors improved classification in 74 patients (23 who died, 51 who survived) but worsened it in 71 patients (4 who died, 67 who survived) with an IDI of 2.4% B. Model with clinical variables and treadmill exercise time compared to model with only clinical variables. Treadmill exercise time with cardiac risk factors improved classification in 70 patients (21 who died, 49 who survived) but worsened it in 68 patients (4 who died, 64 who survived) with an IDI of 1.9% C. Model with clinical variables and treadmill exercise time compared to model with clinical variables and Peak VO₂ improved classification in 43 patients (6 who died, 37 who survived) but worsened it in 46 patients (8 who died, 38 who survived) with an IDI of -0.4% which was not statistically significant

Table 1
 Baseline Characteristics According to Sex-Specific Quartiles for Treadmill Exercise Time

	Quartile 1 (N=553)	Quartile 2 (N=588)	Quartile 3 (N=524)	Quartile 4 (N=566)	P
Treadmill Time: Women (N=602) (range,sec)	21–315	317–478	480–600	603–1320	
Treadmill Time: Men (N=1629) (range,sec)	35–357	360–480	484–657	660–1415	
Age, (yrs)	57 +/- 10	56 +/- 10	54 +/- 11	49 +/- 11	<0.001
Body mass index, (kg/m ²)	29 +/- 6	28 +/- 6	29 +/- 6	28 +/- 5	0.20
Current smokers, n (%)	112 (20)	117 (20)	108 (21)	122 (22)	0.91
Diabetes: insulin treated, n (%)	74 (13)	78 (13)	42 (8)	21 (4)	<0.001
Diabetes: not insulin treated, n (%)	111 (20)	107 (18)	84 (16)	48 (8)	<0.001
Coronary artery disease, n (%)	267 (48)	278 (47)	210 (40)	151 (27)	<0.001
Previous MI, n (%)	68 (12)	63 (11)	76 (15)	72 (13)	0.30
Previous CABG, n (%)	187 (34)	192 (33)	131 (25)	84 (15)	<0.001
Previous PCI, n (%)	135 (24)	147 (25)	123 (23)	71 (13)	<0.001
Implantable cardioverter-defibrillator, n (%)	209 (38)	180 (31)	159 (30)	99 (17)	<0.001
Pacemaker, n (%)	160 (29)	149 (25)	117 (22)	76 (13)	<0.001
Medication use, n (%)					
β-Blocker	349 (63)	357 (61)	345 (66)	378 (67)	0.13
Ace inhibitor	378 (68)	443 (75)	422 (81)	468 (83)	<0.001
Angiotensin receptor blocker	86 (16)	92 (16)	56 (11)	56 (10)	0.003
Potassium sparing diuretics	181 (33)	186 (32)	152 (29)	130 (23)	0.001
Antiarrhythmic	159 (29)	152 (26)	117 (22)	81 (14)	<0.001
Anticoagulation	253 (46)	255 (43)	219 (42)	172 (30)	<0.001
Aspirin	261 (47)	270 (46)	251 (48)	256 (45)	0.81
Digoxin	376 (68)	425 (72)	361 (69)	408 (72)	0.28
Nitrates	212 (38)	223 (38)	162 (31)	142 (25)	<0.001
Hydralazine	44 (8)	40 (7)	29 (6)	23 (4)	0.041
Loop diuretics	502 (91)	530 (90)	447 (85)	401 (71)	<0.001
Thiazide diuretics	90 (16)	98 (17)	53 (10)	38 (7)	<0.001
Statin	210 (38)	266 (45)	187 (36)	187 (33)	<0.001

	Quartiles 1 (N=553)	Quartile 2 (N=588)	Quartile 3 (N = 524)	Quartile 4 (N = 566)	P
Calcium channel blocker:					
Non-dihydropyridine	6 (1)	2 (0)	4 (1)	4 (1)	0.52
Calcium channel blocker: dihydropyridine	24 (4)	33 (6)	31 (6)	11 (2)	0.005
Resting heart rate, (beats/min)	80 +/- 14	78 +/- 14	75 +/- 14	72 +/- 14	<0.001
Resting systolic blood pressure, (mm Hg)	108 +/- 18	110 +/- 18	111 +/- 18	112 +/- 17	<0.001
Left ventricular ejection fraction, (%)	19 +/- 8	20 +/- 7	21 +/- 8	21 +/- 7	<0.001
Peak oxygen consumption, (ml/kg/min)	11 +/- 3	14 +/- 2	17 +/- 3	22 +/- 4	<0.001
Peak respiratory exchange ratio	1.03 +/- 0.16	1.07 +/- 0.11	1.10 +/- 0.09	1.11 +/- 0.07	<0.001
Serum Sodium (mmol/L)	139 +/- 4	139 +/- 3	140 +/- 3	140 +/- 3	<0.001
Creatinine Clearance (ml/min)	78 +/- 37	82 +/- 41	94 +/- 43	110 +/- 44	<0.001
Serum BUN (mg/dL)	31 +/- 16	28 +/- 13	23 +/- 10	19 +/- 7	<0.001
Serum hemoglobin (g/dL)	13 +/- 2	13 +/- 1	14 +/- 1	14 +/- 1	<0.001
Serum glucose (mg/dL)	112 +/- 44	115 +/- 46	111 +/- 47	100 +/- 30	<0.001

N= total number of women and men per quartile

Table 2

Treadmill Exercise Time and Outcome: Cox Proportional Hazards Analyses

Model: <i>Death</i>	Hazard Ratio (95% CI)	P
<u>Treadmill exercise time as a continuous variable[*]</u>		
Unadjusted	2.23 (2.00 to 2.49)	<.0001
Adjusted for age and sex	2.23 (1.98 to 2.50)	<.0001
Adjusted for age, sex, history of CAD, and peak VO ₂	1.54 (1.23 to 1.92)	.0005
Multivariable adjusted	1.40 (1.12 to 1.76)	.004
<u>Treadmill exercise time as a dichotomous variable[^]</u>		
Unadjusted	4.91 (3.86 to 6.24)	<.0001
Multivariable adjusted	1.70 (1.05 to 2.75)	.03

* Comparisons are between the 25th percentile (317 seconds in women, 360 seconds in men) and 75th percentile (600 seconds in women, 657 seconds in men)

[^] Comparisons between quartile 1 (<316 seconds in women, <358 seconds in men) and quartile 4 (>602 seconds in women, >659 seconds in men)

Table 3

Out of Bag (OOB) Concordance Index Values For Cox Regression Models

	All-cause death	
	OOB Concordance Index (95% Confidence Interval)	
Clinical variables only	0.672	(0.646 to 0.698)
Peak VO ₂ only	0.675	(0.662 to 0.688)
Time only	0.666	(0.651 to 0.681)
Peak VO ₂ + Time only	0.679	(0.666 to 0.692)
Clinical variables + Peak VO ₂	0.699	(0.673 to 0.725)
Clinical variables + Time	0.697	(0.671 to 0.723)
Clinical variables + Peak VO ₂ + Time	0.698	(0.670 to 0.726)