Validation of a Cardiopulmonary Exercise Test Score in Heart Failure

Jonathan Myers, PhD; Ricardo Oliveira, PhD; Frederick Dewey, MD; Ross Arena, PhD, PT; Marco Guazzi, MD, PhD; Paul Chase, Med; Daniel Bensimhon, MD; Mary Ann Peberdy, MD; Euan Ashley, MD; Erin West, MS; Lawrence P. Cahalin, PhD, PT; Daniel E. Forman, MD

Background—Cardiopulmonary exercise test (CPX) responses are strong predictors of outcomes in patients with heart failure. We recently developed a CPX score that integrated the additive prognostic information from CPX. The purpose of this study was to validate the score in a larger, independent sample of patients.

Methods and Results—A total of 2625 patients with heart failure underwent CPX and were followed for cardiovascular (CV) mortality and major CV events (death, transplantation, left ventricular assist device implantation). Net reclassification improvement (NRI) for the score and each of its components were determined at 3 years. The VE/VCO2 slope was the strongest predictor of risk and was attributed a relative weight of 7, with weighted scores for abnormal heart rate recovery, oxygen uptake efficiency slope, end-tidal CO2 pressure, and peak VO2 having scores of 5, 3, 3, and 2, respectively. A summed score of >15 was associated with an annual mortality rate of 12.2% and a relative risk >9 for total events, whereas a score of <5 was associated with an annual mortality rate of 1.2%. The composite score was the most accurate predictor of CV events among all CPX responses considered (C indexes, 0.70 for CV mortality and 0.72 for the composite outcome). Each component of the score provided significant NRI compared with peak VO2 (category-free NRI, 0.61–0.77), and the score provided significant NRI above clinical risk factors for both CV events and mortality (NRI, 0.63 and 0.65 for CPX score compared with clinical variables alone).

Conclusions—These results validate the application of a simple, integrated multivariable score based on readily available CPX responses. (Circ Heart Fail. 2013;6:211-218.)

Key Words: epidemiology ■ exercise physiology ■ exercise testing ■ heart failure ■ oxygen consumption

Recent advances in therapy have resulted in a reduction in mortality for most forms of cardiovascular disease. However, success in treating other forms of cardiovascular disease along with aging of the population has resulted in an increase in the prevalence of chronic heart failure (HF).1,2 HF is now the leading cause of hospitalization among those aged >65 years, accounting for ≈20% of hospital admissions in this group.3 Therefore, a great deal of effort in recent years has been directed toward evaluation techniques designed to optimally stratify risk in these patients. A hallmark symptom of HF is exercise intolerance, typically evidenced by excessive shortness of breath, fatigue, or both. During the past 2 decades, the cardiopulmonary exercise test (CPX) has become an important procedure for quantifying the degree of exercise intolerance. Numerous studies have demonstrated that CPX responses have more recently been demonstrated to provide clinically significant and independent information for estimating prognosis in patients with HF.4-6

Clinical Perspective on p 218

There remains debate regarding the optimal application of CPX variables for estimating risk for mortality, hospitalization, or other outcomes in patients with HF. Similar to many other clinical tools, researchers have tended to take a binary approach when applying the CPX for this purpose. For example, a peak VO2 achieved ≤14 mL·kg⁻¹·min⁻¹ has been widely applied to define patients with HF at high risk for adverse events.4,7 More recently, there has been a growing awareness of the additional benefit of applying more complex statistical techniques and multivariate scores to predict risk in patients with cardiovascular disease,4-10 and HF specifically.5,11 The advantage of these approaches is that they permit the quantification of risk across the spectrum of abnormal responses10 and have been

Received February 28, 2012; accepted January 15, 2013.
From the Division of Cardiology, VA Palo Alto Healthcare System, Palo Alto, CA (J.M.); Cardiovascular Medicine, Stanford University, Palo Alto, CA (J.M., F.D., E.A.); Rio de Janeiro State University, Rio de Janeiro, Brazil (R.O.); Physical Therapy Program, Department of Orthopedics and Rehabilitation, and Division of Cardiology, Department of Internal Medicine, University of New Mexico School of Medicine, Albuquerque, NM (R.A.); Division of Cardiology, University of Milano, Milan, Italy (M.G.); Lebauer Cardiovascular Research Foundation, Greensboro, NC (P.C., D.B.); Department of Internal Medicine, Virginia Commonwealth University, Richmond, VA (M.A.P.); Division of Cardiovascular Medicine, Brigham and Women’s Hospital, Boston, MA (E.W., D.E.F.); and Department of Physical Therapy, Leonard M. Miller School of Medicine, University of Miami, Miami, FL (L.P.C.).
Correspondence to Jonathan Myers, PhD, VA Palo Alto Health Care System, Cardiology Division (111C), 3801 Miranda Ave, Palo Alto, CA 94304.
E-mail Drj993@aol.com
© 2013 American Heart Association, Inc.

Circ Heart Fail is available at http://circheartfailure.ahajournals.org DOI: 10.1161/CIRCHEARTFAILURE.112.000073
demonstrated to predict risk more accurately.6,10-12 These approaches have long been recommended for the standard exercise test to assist with the diagnosis of coronary artery disease8,9,13,14 and have recently been applied to CPX for estimating prognosis in patients with HF.6,11

We recently developed a CPX score using a summation of readily available responses that improved the prognostic utility of the test.6 CPX responses recently shown to be strong and independent predictors of outcomes in patients with HF provided incremental, progressive, and independent information to the prediction of adverse outcomes. However, the sample used was relatively small, and a validation cohort for the score was not available. In addition, the association between a given risk marker and outcomes, despite generating a significant hazard, does not necessarily result in a higher reclassification of risk.5,16 Recently, statistical tests such as the net reclassification improvement (NRI) have been recommended to better quantify the ability of a measure to discriminate risk. The NRI improves on more standard indices of predictive modeling, such as the area under the receiver operating characteristic curve in that it more directly and incrementally evaluates the ability of new risk markers to classify subjects into higher or lower categories of risk.15,17 In the current study, we sought to: (1) validate a CPX score developed previously6 in a larger, independent sample of patients with HF; and (2) determine the contributions of individual and combined components of the CPX to enhance risk classification in patients with HF.

Methods

This study was performed as part of an HF consortium; a multicenter, retrospective analysis including patients with HF from the exercise laboratories at the VA Palo Alto Health Care System and Stanford University, Palo Alto, CA; San Paolo Hospital, Milan, Italy; Virginia Commonwealth University, Richmond, VA; Brigham and Women’s Hospital, Boston, MA; and the LeBauer Cardiovascular Research Foundation, Greensboro, NC. A total of 2625 patients with chronic HF, tested between 1993 and 2010, were included. The sample included 1974 men and 651 women, with a mean age of 56±14 years. Eighty-nine percent of the subjects were independent from the original sample from which the score was developed. Inclusion criteria consisted of a diagnosis of HF6 and evidence of left ventricular systolic dysfunction (ejection fraction [EF] <40%) or HF with preserved EF by 2-dimensional echocardiography obtained within 1 month of exercise testing. HF with preserved systolic function was considered to be present if the EF was normal (>45%) and the subject had a history of decompensated HF. Subjects received routine follow-up care at the 5 institutions included in the study. All subjects were stable and receiving optimal medical therapy at the time of testing. The subjects completed a written informed consent, and institutional review board approval was obtained at each institution.

CPX Procedure and Data Collection

Symptom-limited CPX was performed on all patients using treadmill or cycle ergometer ramping protocols.19 A treadmill was used for testing in the American centers, whereas a cycle ergometer was used in the European center. We previously observed that optimal peak VO2 and VE/VCO2 slope threshold values for estimating prognosis were similar irrespective of mode of exercise in patients with HF.20 Ventilatory expired gas analysis was performed using a metabolic cart at all 5 centers (Medgraphics CPX-D or ULTIMA PFX, Minneapolis, MN; Orca Diagnostics, Santa Barbara, CA; Parvo Medics TrueOne 2400, Sandy, UT; or CareFusion Oxygen Pro, San Diego, CA). Before each test, the equipment was calibrated in standard fashion using reference gases. A standard 12-lead ECG was obtained at rest, each minute during exercise, and for at least 5 minutes during the recovery phase; blood pressure was measured using a standard cuff sphygmomanometer.

Minute ventilation (VE, body temperature and pressure, saturated [BTPS]), oxygen uptake (VO2, standard temperature and pressure, dry [STPD]), carbon dioxide production (VCO2, STPD), and other CPX variables were acquired breath-by-breath and averaged over 10- or 15-second intervals. VE and VCO2 responses throughout exercise were used to calculate the VE/VCO2 slope via least squares linear regression (y=mx+b, where m=slope). Previous work by our group and others has shown this method of calculating the VE/VCO2 slope to be optimal for estimating prognosis.6,12 The oxygen uptake efficiency slope (OUES) was calculated using [(VO2 (L/min) = m (log10 VE)+b, where m=OUES).21 HRR was defined as maximal heart rate minus heart rate at 1 minute in recovery.22 Resting end-tidal CO2 pressure (PetCO2) was derived from the average of a 2-minute sitting resting period before the test.23

End Points

The primary end point was cardiac-related mortality. A second composite end point including major cardiac events was also studied; this included cardiac transplantation, left ventricular assist device (LVAD) implantation, and cardiac-related death. Subjects were followed for major cardiac-related events for 3 years after their exercise test using the Social Security Death Index and hospital and outpatient medical chart review. Follow-up was performed by the HF program at each respective institution, providing a high likelihood that all major events were captured. Individuals conducting the CPX were not involved in decisions regarding cause of death or heart transplant/LVAD implantation.

Statistical Analysis

NCSS (Kayesville, UT) software and the Design and Hmisc libraries in S-Plus 7.0 and R (Seattle, WA) were used for all statistical analyses. Unpaired t tests were used for comparisons of continuous variables, and χ2 tests were used to compare categorical variables between those who experienced a cardiac event and those who did not. Receiver operating characteristic curve analysis was used to define optimal threshold values for each CPX response. Z tests were used to compare the areas under the receiver operating characteristic curves for CPX responses. Cox proportional hazards analysis was used to determine age-adjusted hazard ratios for the 5 CPX variables included in the model, each expressed dichotomously using the threshold value. Optimal thresholds for each of the CPX variables were as follows: VE/VCO2 slope (≥3.4) abnormal HRR (≥6 beats at 1 minute), OUES (≤1.4), PetCO2 (<33 mm Hg), and peak VO2 (≤14 mL·kg−1·min−1). Each variable was assigned a weight according to the hazard ratios and summed to calculate the composite score. Proportional hazard assumptions were confirmed for each variable using the log [−log (survival function)] plot.

Kaplan–Meier analysis was used to determine overall and cardiovascular event-free survival characteristics for the summed score classifications 0 to 5, 6 to 10, 11 to 15, and >15. This analysis was repeated in 2 prespecified subgroups, which comprised subjects with left ventricular EF (LVEF) <30% and subjects with LVEF ≥30%. The log-rank test was used to determine statistical significance of the Kaplan–Meier analyses. Multivariable Cox proportional hazards analysis adjusted for age, sex, body mass index, EF, and HF pathogenesis was then used to calculate hazard ratios for each summed score classification group.

The predictive accuracies of each of the CPX responses were determined using both the right-censored concordance index (C index) validated with 200 bootstrap samples and the Akaike Information Criterion method.24 The predictive accuracy of the summed score was then evaluated via similar analyses in 4 prespecified subgroups: subjects with ischemic cardiomyopathy and nonischemic cardiomyopathy, and subjects with LVEF <30% and LVEF ≥30%. To further evaluate the reclassification characteristics of individual components of the CPX score, as well as the composite CPX score in comparison with standard clinical risk factors (age, systolic blood pressure, HF pathogenesis, body mass index, diuretic use, and LVEF), we
calculated the category-free NRI index, modified for right-censored survival data according to the methods proposed by Pencina et al.\textsuperscript{27} The NRI was calculated for both cardiovascular-related mortality and major cardiovascular events. We corrected for overoptimism using 1000 bootstrap replicates and reported the median results and bootstrap estimated 95% confidence intervals (CIs).

**Results**

**Baseline Characteristics of the Study Population and Development of the Summed Score**

The study sample comprised 1974 men and 651 women with HF; 35% had an ischemic pathogenesis. The mean age of the cohort was 56±14 years, and the mean body mass index was 28.7±6.0 kg/m\(^2\). Subjects who died from cardiac causes were older and had a lower EF compared with subjects with no events (Table 1).

Among CPX variables, peak VO\(_2\) (18.6±8.5 vs 14.1±5.4 mL·kg\(^{-1}\)·min\(^{-1}\)), peak heart rate, HRR, OUES, and PetCO\(_2\) were higher among those with no events. Conversely, resting heart rate, the VE/VCO\(_2\) slope and the CPX weighted summed score were lower among in the no event group versus those who died from cardiac causes. Peak VO\(_2\) was lower, whereas the VE/VCO\(_2\) slope and the weighted summed score were higher in patients who had a secondary outcome (LVAD or transplantation) versus both the no event and cardiac mortality groups.

There were 412 total adverse events (290 deaths, 79 transplantations, and 43 LVAD implantations) during the mean 2.4±2.5 year follow-up. The weighted scores for abnormal CPX responses were derived from proportional hazards analysis and replicated from the previous score:\textsuperscript{6} weighted scores of 7, 5, 3, and 2 were applied for the VE/VCO\(_2\) slope, HRR, OUES, PetCO\(_2\), and peak VO\(_2\), respectively. When only those patients taking β-blockers were studied, the relative weights were similar, with the exception that there was a lower weight for HRR (weight=2).

**Predictors of Adverse Events**

Age-adjusted univariate predictors of cardiovascular mortality, secondary events, and total events are presented in Table 2. Each of the CPX responses in the score was significantly associated with each of the outcomes, with an abnormal VE/VCO\(_2\) slope generating the highest risk (hazard ratios, 3.2 [95% CI, 2.5–4.3]; 8.3 [95% CI, 5.4–12.9]; and 4.3 [95% CI, 3.5–5.3]; all P<0.001) for cardiovascular mortality, secondary events, and total events, respectively). With a weighted summed score of 0 to 5 as the reference group, risks for all event categories were significantly higher as the weighted summed score categories increased from 6 to 10, 11 to 15, and >15. For total events, a score of >15 was associated with a hazard ratio of >9. These results were similar among patients with preserved and reduced EF. Each CPX response and the composite score also significantly predicted risk when expressed as continuous variables.

**Relationships Between Summed Score and Outcomes**

Overall mortality and the composite event-free Kaplan–Meier survival estimates according to summed score classifications are presented in Figures 1 and 2, respectively. There were significant
stepwise increases in both mortality and composite outcome rates associated with increasing weighted summed scores. The estimated 1-year death rate was 12.2% for subjects with a summed score of >15 and only 1.2% for subjects with a summed score of <5. Similarly, whereas subjects with summed scores of >15 had estimated 1-year rates of death, transplantation, or LV AD of 17%, subjects with summed scores of <5 had rates of such events at 1 year of 2.8%. This stepwise increase in risk persisted in Kaplan–Meier subgroup analyses for both subjects with LVEF >30% and subjects with LVEF ≤30%, although subjects with LVEF >30% had lower overall event rates.

Predictive Accuracy of CPX Variables and Summed Score

The predictive accuracy of CPX responses and the weighted summed score for mortality and major events are presented in Table 3. The VE/VCO₂ slope was the most accurate predictor of outcomes among individual CPX variables (C index, 0.70 for major events), followed by peak VO₂, OUES, PetCO₂, and HRR. The summed risk score was a more accurate predictor of outcomes than any individual CPX variable (C indexes, 0.70 for cardiac mortality and 0.72 for major events, respectively). The predictive accuracy of the summed score for mortality and the composite outcome was similar in subjects with ischemic and nonischemic cardiomyopathy.

Table 4 presents age-adjusted Akaike Information Criterion weights for each individual CPX response and the summed score. The summed score had the highest predictive value (0.73, indicating a 73% probability of being the strongest model). The score remained the most powerful after adjustment for β-blocker use and after applying different cut points for high risk.

Table 2. Age-adjusted Univariate Proportional Hazards Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cardiac Mortality (n=224)</th>
<th>LVAD/Transplant (n=121)</th>
<th>Total Events (n=412)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR 95% CI P Value</td>
<td>HR 95% CI P Value</td>
<td>HR 95% CI P Value</td>
</tr>
<tr>
<td>Age, y</td>
<td>1.37 1.20–1.57 &lt;0.0001</td>
<td>1.46 1.21–1.77 &lt;0.0001</td>
<td>1.10 0.98–1.21 0.11</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>1.11 0.97–1.27 0.14</td>
<td>1.09 0.90–1.33 0.39</td>
<td>1.11 1.00–1.25 0.06</td>
</tr>
<tr>
<td>Ejection fraction, %</td>
<td>1.55 1.33–1.81 &lt;0.0001</td>
<td>3.33 2.56–4.33 &lt;0.0001</td>
<td>1.95 1.71–2.23 &lt;0.0001</td>
</tr>
<tr>
<td>NYHA Class</td>
<td>1.92 1.39–2.66 &lt;0.0001</td>
<td>3.90 2.65–5.75 &lt;0.0001</td>
<td>2.53 1.99–3.21 &lt;0.0001</td>
</tr>
<tr>
<td>Pathogenesis (ischemic)</td>
<td>1.01 1.36–2.34 &lt;0.0001</td>
<td>1.98 1.34–2.95 0.0007</td>
<td>1.72 1.40–2.10 &lt;0.0001</td>
</tr>
<tr>
<td>Peak VO₂≤14 mL·kg⁻¹·min⁻¹</td>
<td>1.98 1.52–2.57 &lt;0.0001</td>
<td>6.07 4.06–9.05 &lt;0.0001</td>
<td>3.10 2.54–3.77 &lt;0.0001</td>
</tr>
<tr>
<td>HRR≤6 beats at 1 min</td>
<td>1.75 1.09–2.79 0.02</td>
<td>2.04 0.99–4.22 0.054</td>
<td>2.14 1.55–2.96 &lt;0.0001</td>
</tr>
<tr>
<td>VE/VCO₂ slope ≥34</td>
<td>3.24 2.47–4.26 &lt;0.0001</td>
<td>8.32 5.36–12.92 &lt;0.0001</td>
<td>4.29 3.48–5.29 &lt;0.0001</td>
</tr>
<tr>
<td>OUES&lt;1.4</td>
<td>2.33 1.56–3.44 &lt;0.0001</td>
<td>5.75 2.79–11.90 &lt;0.0001</td>
<td>2.87 2.08–3.94 &lt;0.0001</td>
</tr>
<tr>
<td>PetCO₂&lt;33 mm Hg</td>
<td>1.77 1.25–2.51 0.001</td>
<td>3.11 1.96–4.95 &lt;0.0001</td>
<td>2.38 1.86–3.04 &lt;0.0001</td>
</tr>
</tbody>
</table>

Weighted summed score

| 0 to 5 (n=1398) | 1 (Reference) | 1 (Reference) | 1 (Reference) |
| 6 to 10 (n=841) | 2.33 1.72–3.16 <0.0001 | 5.21 3.26–8.33 <0.0001 | 2.74 2.16–3.48 <0.0001 |
| 10 to 15 (n=335) | 3.26 2.28–4.67 <0.0001 | 7.16 4.20–12.21 <0.0001 | 4.6 3.55–5.98 <0.0001 |
| >15 (n=38)     | 4.31 1.87–9.9 0.0005 | 12.4 4.72–32.57 <0.0001 | 9.25 5.75–14.88 <0.0001 |

Age, BMI, and ejection fraction are increments using SD. BMI indicates body mass index; CI, confidence interval; HR, hazard ratios; HRR, heart rate recovery at 1 min; LVAD, left ventricular assist device; NYHA, New York Heart Association; OUES, oxygen uptake efficiency slope; and PetCO₂, end-tidal carbon dioxide pressure.

*Includes LVAD, transplant, and all causes of death.
Classification of Risk

Table 5 presents category-free NRI indexes for major cardiovascular events at 36 months for individual components of the CPX score. The VE/VCO₂ slope, OUES, end-tidal PetCO₂, and HRR all provided significant overall incremental risk reclassification. The risk reclassification improvement provided by HRR was specific to individuals without cardiovascular events; all other components of the CPX score provided significant risk reclassification improvement both for subjects with cardiovascular events and those without cardiovascular events at 36 months. Table 6 presents category-free NRI for the CPX score in comparison with standard clinical risk factors (age, systolic blood pressure, HF pathogenesis, body mass index, diuretic use, and LVEF). The CPX score provided significant NRI for both cardiovascular-related mortality (NRI=0.65; 95% CI, 0.61–0.69) and for cardiovascular events (NRI=0.63; 95% CI, 0.59–0.68) at 36 months.

Discussion

The current results provide a measure of validation for previously developed CPX score⁶ in a group of patients with HF that is significantly larger than the sample from which the score was developed. The multivariate score used common and easily derived CPX responses, and its application improved the classification risk for adverse events in patients with HF. The estimation of risk was incremental, with each component of the score adding progressively and independently to the prediction of outcomes. The composite CPX score was the most accurate predictor of outcomes among all the CPX responses considered, and accurately predicted risk for adverse events among patients with both ischemic and nonischemic cardiomyopathy and among both patients with LVEF >30% and ≤30%. The current data thus represent what is commonly termed a validation set or a measure of cross-validation for the original score.²⁸ These findings further refine the application of CPX for the estimation of risk in patients with HF and may help to optimize the clinical decision-making process when evaluating these patients.

We used a novel index of risk classification, the NRI,¹⁵–¹⁷ to provide better insight into the individual components of CPX.

Table 3. Predictive Accuracy of CPX Parameters and Composite Risk Score

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cardiac Mortality</th>
<th>Major Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak VO₂≤14 mL·kg⁻¹·min⁻¹</td>
<td>0.61</td>
<td>0.67</td>
</tr>
<tr>
<td>HRR≤6 beats at 1 min</td>
<td>0.61</td>
<td>0.60</td>
</tr>
<tr>
<td>VE/VCO₂ slope ≥34</td>
<td>0.66</td>
<td>0.70</td>
</tr>
<tr>
<td>OUES≤1.4</td>
<td>0.62</td>
<td>0.65</td>
</tr>
<tr>
<td>PetCO₂&lt;33 mmHg</td>
<td>0.58</td>
<td>0.61</td>
</tr>
<tr>
<td>Composite risk score</td>
<td>0.70</td>
<td>0.72</td>
</tr>
</tbody>
</table>

CPX indicates cardiopulmonary exercise test; HRR, heart rate recovery at 1 min; OUES, oxygen uptake efficiency slope; and PetCO₂, end-tidal carbon dioxide pressure.
that are known to predict risk.4,6 The NRI reflects clinically meaningful improvement in risk classification achieved with each component of the score; it is calculated as the net change in risk among subjects after the addition of each marker to the baseline model (in the current case, standard clinical variables). Our results extend previous findings by demonstrating that each CPX response included in an integrated score (the VE/VCO₂ slope, peak VO₂, OUES, resting PetCO₂, and HRR) added significant incremental improvement in risk reclassification. For example, adding individual components of the score to peak VO₂ improved classification of subjects at risk (up or down) by ≈70% (Table 5); similarly, adding the integrated CPX score to clinical variables improved the classification of subjects at risk by >60% (Table 6). Thus, the addition of the score to a simple model of clinical risk factors significantly improved risk classification. In addition, these findings suggest that the individual components of the score reflect separate maladaptive pathways in HF and that each contributes to adverse event risk.

Although peak VO₂ has been the most widely used CPX variable to predict risk, reliance on any single factor or statistic is generally known to have limited accuracy.15,29 For example, multivariate scores have long been recommended to enhance the diagnostic and prognostic accuracy of the standard exercise test and have been recommended in exercise testing guidelines.14 This is in part because of a growing awareness of the need to apply statistical techniques to develop evidence-based multivariable models for improving clinical decision making.10,12 The CPX score described herein is consistent with this approach, because it provides quantification of risk across the spectrum of abnormal responses. The performance of the score was similar to the original analysis6 both in terms of its predictive accuracy and overall risk using cumulative scores (Tables 3 and 4; Figures 1 and 2). Patients with a summed score of >15 had a >4-fold risk for cardiovascular mortality, a >12-fold risk for secondary events, and a >9-fold risk for total events. This is contrasted by the risk associated with, for example, an impaired peak VO₂ alone, which had an age-adjusted hazard ratio of 2.0, and underscores the advantages of applying a multivariable approach as opposed to the commonly applied binary method. The CPX score yielded a 73% probability that the model was superior, as compared with the negligible probability when using any 1 of the variables alone, or even when compared with the combination of peak VO₂ and the VE/VCO₂ slope (Table 4).

### Table 5. Net Reclassification Improvement for Major Events at 36 Months According to Components of CPX Score

<table>
<thead>
<tr>
<th>Model</th>
<th>Event-free Survival Rate (All Subjects)</th>
<th>Event-free Survival Rate (Subjects Reclassified Up)</th>
<th>Event-free Survival Rate (Subjects Reclassified Down)</th>
<th>NRI, Subjects With Events</th>
<th>NRI, Subjects Without Events</th>
<th>NRI, All Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak VO₂+VE/VCO₂ slope*</td>
<td>0.87</td>
<td>0.77 (185)</td>
<td>0.92 (293)</td>
<td>0.31 (0.27 to 0.34)</td>
<td>0.30 (0.29 to 0.31)</td>
<td>0.61 (0.57 to 0.64)</td>
</tr>
<tr>
<td>Peak VO₂+VE/VCO₂ slope+OUES†</td>
<td>0.69 (131)</td>
<td>0.93 (347)</td>
<td>0.22 (0.18 to 0.25)</td>
<td>0.55 (0.54 to 0.56)</td>
<td>0.77 (0.73 to 0.80)</td>
<td></td>
</tr>
<tr>
<td>Peak VO₂+VE/VCO₂ slope+OUES+PetCO₂†</td>
<td>0.71 (188)</td>
<td>0.93 (290)</td>
<td>0.37 (0.33 to 0.40)</td>
<td>0.29 (0.28 to 0.30)</td>
<td>0.66 (0.62 to 0.69)</td>
<td></td>
</tr>
<tr>
<td>Peak VO₂+VE/VCO₂ slope+OUES+PetCO₂+HRR†</td>
<td>0.63 (79)</td>
<td>0.92 (399)</td>
<td>−0.02 (−0.05 to 0.01)</td>
<td>0.76 (0.76 to 0.78)</td>
<td>0.74 (0.71 to 0.78)</td>
<td></td>
</tr>
<tr>
<td>CPX indicates cardiopulmonary exercise test; HRR, heart rate recovery at 1 min; OUES, oxygen uptake efficiency slope; PetCO₂, end-tidal carbon dioxide pressure; and NRI, net reclassification improvement.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Compared with peak VO₂ only.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>†Compared with model described in preceding row.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‡Numbers in parentheses represent numbers of subjects reclassified up or down.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6. Net Reclassification Improvement for Major Cardiovascular Events and Cardiovascular-Related Mortality at 36 Months According to CPX Score

<table>
<thead>
<tr>
<th>Major cardiovascular events</th>
<th>Event-free Survival Rate (All Subjects)</th>
<th>Event-free Survival Rate (Subjects Reclassified Up)</th>
<th>Event-free Survival Rate (Subjects Reclassified Down)</th>
<th>NRI, Subjects With Events</th>
<th>NRI, Subjects Without Events</th>
<th>NRI, All Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical variables</td>
<td>0.80</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Clinical variables+CPX score*</td>
<td>0.69 (585)</td>
<td>0.88 (826)</td>
<td>0.36 (0.32–0.39)</td>
<td>0.27 (0.26–0.29)</td>
<td>0.63 (0.59–0.68)</td>
<td></td>
</tr>
</tbody>
</table>

**Cardiovascular-related mortality**

| Clinical variables         | 0.88                                   | ...                                                | ...                                                  | ...                       | ...                        | ...               |
| Clinical variables+CPX score* | 0.80 (568)                             | 0.93 (842)                                         | 0.36 (0.33–0.40)                                     | 0.29 (0.28–0.31)          | 0.65 (0.61–0.69)          |                   |

*CPX indicates cardiopulmonary exercise test; and NRI, net reclassification improvement. Compared with model including clinical variables only: age, systolic blood pressure, heart failure etiology, body mass index, diuretic use, and left ventricular ejection fraction. †Numbers in parentheses represent numbers of subjects reclassified up or down.
In recent years, our group6,20,21,24,25 and others5,22,30 have demonstrated significant and independent prognostic value for each of the components of the score. One explanation for the strong and incremental prognostic power we observed is that the components of the score reflect different pathologies that are characteristic of HF. Although peak VO₂ has long been recognized as an important prognostic marker in patients with HF, indices of ventilatory inefficiency have more recently been demonstrated to be important components of the risk paradigm,4,5 and their application is now advocated in guidelines on HF management.31,32 In particular, the VE/VCO₂ slope has been widely studied and has been shown to be a more powerful predictor of risk than peak VO₂,4,6,31 which is consistent with the current study (Tables 2 and 3). Abnormalities in ventilatory efficiency have been demonstrated to reflect ventilation/perfusion mismatching in the lungs (related, in part, to an impaired cardiac output response to exercise), early lactate accumulation, and abnormalities in respiratory control.4,5,31,32 Abnormal HRR is associated with autonomic dysfunction that is common in HF (reflected by impaired vagal reactivation) and has been shown to provide prognostic power independent of peak VO₂ and the VE/VCO₂ slope.24

The present results extend previous findings from Aaronson et al11 who used peak VO₂ along with several noninvasive clinical markers in a multivariate model to predict event-free survival in patients with HF (termed the Heart Failure Survival Score [HFSS]). Application of 7 noninvasive variables identified low- and high-risk groups with 93% and 43% 1-year event-free survival, respectively. The HFSS has outperformed peak VO₂ alone in both US and European populations of patients with HF;11,33,34 Although the HFSS has been validated33 and widely used, it included only peak VO₂ from CPX. Numerous recent studies have also incorporated indices of ventilatory inefficiency in addition to peak VO₂ to predict prognosis in HF.4,6 All of these studies have demonstrated improved risk stratification by documenting inefficient ventilation in addition to impaired peak VO₂. However, we are unaware of other multivariate scores focusing specifically on CPX responses.

Limitations
By design, the CPX score focuses primarily on the ventilatory gas exchange response to exercise and does not include other clinical markers of risk in HF. There are other CPX responses that predict risk, particularly respiratory breathing,4,31 which were not included in the score. In addition, although the score was compared with simple clinical variables, it should be noted that there are many other variables and biomarkers that have been used to define risk in HF. A more complex score including some of these markers may provide better precision for estimating risk. Some patients had LVAD or transplantation as their end point, and given patients’ CPX responses likely influenced the decision to have these procedures, raising a potential bias. Finally, the sample was 75% men, and the results may not be as applicable to women.

Summary
The current results extend the many recent studies demonstrating the strong prognostic value of CPX. Our findings validate a composite CPX score6 for predicting risk of adverse events in patients with HF; individual components of the score improved reclassification of risk for mortality and adverse events. The simple summation of easily derived responses from CPX can be applied to more accurately estimate risk in patients with HF.

Disclosures
None.

References


**CLINICAL PERSPECTIVE**

The cardiopulmonary exercise test (CPX) has been widely used in recent years to stratify risk in patients with heart failure. However, the optimal method of applying CPX responses remains a topic of debate. We recently developed a multivariable CPX score that integrated the additive prognostic information from 5 CPX responses. The purpose of this study was to validate the score in a larger, independent sample of patients with heart failure. We studied 2625 adults with heart failure who underwent CPX and were followed for a mean of 29±30 months. The score was derived by weighting the age-adjusted prognostic power of 5 CPX variables using a summary of point-based risk scores. The VE/VCO2 slope (≥34) was attributed a relative weight of 7, with weighted scores for abnormal heart rate recovery at 1 minute, oxygen uptake efficiency slope, resting end-tidal CO2 pressure, and peak VO2 having scores of 5, 3, 3, and 2, respectively. A summed score of >15 was associated with an annual mortality rate of 12.2% and a relative risk of 9.2 for total events, whereas a score of <5 was associated with an annual mortality of 1.2%. The composite score was the most accurate predictor of cardiovascular events among all CPX responses considered. Each individual component of the score provided significant net reclassification improvement compared with peak VO2 and the score provided significant net reclassification improvement above clinical risk factors. These results validate the application of a simple, integrated multivariable CPX score; the score markedly improved risk classification when added to clinical variables, peak VO2, and other CPX responses.
Validation of a Cardiopulmonary Exercise Test Score in Heart Failure
Jonathan Myers, Ricardo Oliveira, Frederick Dewey, Ross Arena, Marco Guazzi, Paul Chase, Daniel Bensimhon, Mary Ann Peberdy, Euan Ashley, Erin West, Lawrence P. Cahalin and Daniel E. Forman

Circ Heart Fail. 2013;6:211-218; originally published online February 7, 2013;
doi: 10.1161/CIRCHEARTFAILURE.112.000073
Circulation: Heart Failure is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-3289. Online ISSN: 1941-3297

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circheartfailure.ahajournals.org/content/6/2/211

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Heart Failure can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Heart Failure is online at:
http://circheartfailure.ahajournals.org//subscriptions/