Association of Functional and Health Status Measures in Heart Failure

JONATHAN MYERS, PhD, NAYMA ZAHEER, MD, SUSAN QUAGLIETTI, RNP, RUPA MADHAVAN, BS, VICTOR FROELICHER, MD, AND PAUL HEIDENREICH, MD

Palo Alto, California

ABSTRACT

Background: A wide variety of instruments have been used to assess the functional capabilities and health status of patients with chronic heart failure (HF), but it is not known how well these tests are correlated with one another, nor which one has the best association with measured exercise capacity.

Methods and Results: Forty-one patients with HF were assessed with commonly used functional, health status, and quality of life measures, including maximal cardiopulmonary exercise testing, the Duke Activity Status Index (DASI), the Veterans Specific Activity Questionnaire (VSAQ), the Kansas City Cardiomyopathy Questionnaire (KCCQ), and 6-minute walk distance. Pretest clinical variables, including age, resting pulmonary function tests (forced expiratory volume in 1 s and forced vital capacity), and ejection fraction (EF) were also considered. The association between performance on these functional tools, clinical variables, and exercise test responses including peak VO₂ and the VO₂ at the ventilatory threshold, was determined. Peak oxygen uptake (VO₂) was significantly related to VO₂ at the ventilatory threshold ($r = 0.76, P < .001$) and estimated METs from treadmill speed and grade ($r = 0.72, P < .001$), but had only a modest association with 6-minute walk performance ($r = 0.49, P < .01$). The functional questionnaires had modest associations with peak VO₂ ($r = 0.37, P < .05$ and $r = 0.26$, NS for the VSAQ and DASI, respectively). Of the components of the KCCQ, peak VO₂ was significantly related only to quality of life score ($r = 0.46, P < .05$). Six-minute walk performance was significantly related to KCCQ physical limitation ($r = 0.53, P < .01$) and clinical summary ($r = 0.44, P < .05$) scores. Among pretest variables, only age and EF were significantly related to peak VO₂ ($r = -0.58$, and 0.46, respectively, $P < .01$). Multivariately, age and KCCQ quality of life scores were the only significant predictors of peak VO₂, accounting for 72% of the variance in peak VO₂.

Conclusion: Commonly used functional measures, symptom tools, and quality of life assessments for patients with HF are poorly correlated with one another and are only modestly associated with exercise test responses. These findings suggest that exercise test responses, non-exercise test estimates of physical function, and quality of life indices reflect different facets of health status in HF and one should not be considered a surrogate for another.

Key Words: Heart failure, Exercise testing, Maximal oxygen uptake, Exercise capacity.

Heart failure (HF) is characterized by exercise intolerance, typically from shortness of breath or fatigue. The mechanisms for exercise limitation in HF have been the topic of a great deal of investigation, and include factors such as a diminished cardiac output response to exercise, skeletal muscle metabolic dysfunction, abnormal vasodilatory capacity, and abnormal distribution of blood flow. The degree of exercise intolerance is commonly used clinically to assess the severity of the disease and efficacy of therapy. Randomized clinical trials frequently use exercise duration or workload on a standardized exercise test as a primary outcome variable when assessing the efficacy of an intervention. The most accurate expression of exercise tolerance is directly measured peak oxygen uptake (peak VO₂), although numerous surrogates for peak VO₂ and other expressions of a patient's functional capabilities have been used, including symptom questionnaires, submaximal walking tests, and various functional classifications.

Peak VO₂ is rarely performed in clinical trials among patients with HF because of the added time, equipment, and
technical expertise required to use ventilatory gas exchange equipment. Advantages of submaximal tests or health status and symptom questionnaires include that they are quick, inexpensive, and safe. Examples of health status and functional measures commonly used include the Kansas City Cardiomyopathy Questionnaire (KCCQ), Veterans Specific Activity Questionnaire (VSAQ), Duke Activity Status Index (DASI), 6-minute walk distance, and New York Heart Association (NYHA) functional class. Although it has been suggested that these measures can be unreliable because they may provide incomplete or misleading information, some of these instruments have also been shown to be sensitive to changes in clinical status. However, it is uncertain how well the commonly used measures of functional status in patients with HF are associated with clinical status measures, quality of life, or directly measured exercise capacity.

In the present study, we evaluated the association between peak VO2, clinical measures, and commonly used symptom and functional tools in patients with HF. Our objectives were: (1) to evaluate the relationships between these functional assessments and to determine which is most strongly associated with peak VO2 and (2) to determine the extent to which these tools could be used to predict peak VO2.

**Methods**

**Patients**

Forty-one male patients with stable HF (mean age 68 ± 12 years) participated in the study. HF was documented by clinical history and diagnosis in outpatient medical records, and an ejection fraction (EF) <40%. Ischemic HF (n = 22) was defined by having a history of myocardial infarction, coronary bypass surgery, or documented angiographic coronary disease. Only patients with stable, compensated HF >30 years of age were included in the study. Written informed consent was obtained from all subjects using a protocol approved by the Stanford Investigational Review Board. Clinical characteristics of the study group are presented in Table 1.

**Exercise Testing**

Symptom limited maximal exercise tests were performed on a treadmill using an individualized ramp protocol. This test individualizes warm-up and peak walking speeds, along with the ramp rate (rate of change in speed and grade) to yield a test duration of approximately 10 minutes. All tests were performed in the morning. Before testing, standard pulmonary function tests were performed, including forced vital capacity (FVC) and forced expiratory volume in one second (FEV1). Ventilatory oxygen uptake was measured using an Orca Diagnostics metabolic system (Santa Barbara, CA). Gas exchange data were acquired breath-by-breath and expressed in 10-s intervals of rolling 30-s averages. Ventilatory oxygen uptake, carbon dioxide production, minute ventilation, and respiratory exchange ratio were calculated online. The percentage of age-predicted normal peak VO2 was determined for each patient using the equation of Wasserman et al. The ventilatory threshold was determined by 2 experienced, independent reviewers using the V-slope method, and confirmed by ventilatory criteria. Estimated peak VO2 was determined from the American College of Sports Medicine equations. A 12-lead electrocardiogram was monitored continuously and recorded every minute. Blood pressure was recorded manually every 2 minutes throughout the test. All subjects were encouraged to provide a maximal effort, and the Borg 6 to 20 perceived exertion scale was used to quantify effort.

**Symptom Questionnaires**

Before undergoing exercise testing, responses to 4 symptom and health status questionnaires were obtained. These included the KCCQ, VSAQ, DASI, and NYHA functional class. The KCCQ is a detailed, disease-specific health status measure that encompasses domains including physical limitations, symptoms, disease severity, change in status over time, self-efficacy, social interference, and quality of life. The DASI is a brief, self-administered questionnaire designed to estimate functional status. The VSAQ was similarly developed by our group and was specifically designed to estimate exercise tolerance before exercise testing to individualize the treadmill protocol for clinical exercise testing. The VSAQ score was age-adjusted. The NYHA Functional Classification is a widely used general (4 categories), provider-derived classification scheme to categorize patients in terms of symptoms associated with daily activities. A description, metric, and ranges for each of the health status measurements is presented in Table 2.

**Six-Minute Walk Test**

The 6-minute walk test was completed after the exercise test after a minimum 30-minute rest period, and performed as described by Guyatt et al. A 100-foot distance was marked in a hospital corridor, and patients were instructed to walk from end to end at their own pace, attempting to cover as great a distance as possible within 6 minutes. The test was supervised by a technician, who provided verbal encouragement at 30-second intervals.
were permitted to stop and rest, and were instructed to continue walking as soon as they felt they were able to do so. Symptoms experienced by the patient were recorded. Distance covered was expressed in meters.

Statistics

Simple univariate linear regression was performed to assess the relationships between peak VO₂ and the various functional and quality of life measures, including 6-minute walk distance, age-adjusted VSAQ score in metabolic equivalents (METS), functional capacity as determined by the DASI, the functional, clinical status, and quality of life domains of the KCCQ (each expressed in 0 to 100 scales), NYHA functional class, VO₂ at the ventilatory threshold, and peak METS estimated from treadmill speed and grade. Using the pretest variables with the strongest univariate associations with peak VO₂ (age, 6-minute walk distance, VSAQ score, FVC, and KCCQ quality of life score), a forward stepwise multiple regression procedure was performed to predict peak VO₂. A second multiple regression procedure was performed to determine predictors of estimated METS achieved (from peak treadmill speed and grade), using the variables with the strongest univariate associations with estimated METS (age, DASI, VSAQ, 6-minute walk distance, and KCCQ physical limitation score).

Results

Exercise test results are listed in Table 3. The mean maximal perceived exertion was 17.6 ± 2.0, and the mean peak respiratory exchange ratio was 1.14 ± 0.23, suggesting that maximal effort was achieved by most patients. The mean maximal heart rate of 124 ± 28 beats/min was lower than that expected for age (82% of predicted), reflecting that many patients were limited by symptoms associated with HF, that heart rate was limited by the effects of β-blockade, or both. Mean maximal oxygen uptake was 16.1 ± 6.2 mL·kg·min⁻¹ (representing 54 ± 22% of age-predicted peak VO₂), and the ventilatory threshold occurred at 76% of peak VO₂.

Mean values for the various functional and health status measures are presented in Table 4. The mean NYHA functional class was 1.7 ± 0.98. METS from the VSAQ (6.6 ± 2.0) were higher than that achieved on the treadmill expressed as measured peak VO₂ (4.5 ± 1.8, P < .01), and METS estimated from peak treadmill speed and grade (5.4 ± 2.2, P < .05). Mean 6-minute walk distance was 328 ± 97 meters.

Association Between Functional Measures and Peak VO₂

Correlation coefficients between peak VO₂ and the various functional measures are presented in Table 5. Both estimated METS and VO₂ at the ventilatory threshold were significantly associated with peak VO₂ (r = 0.72 and 0.76, P < .001 and P < .001, respectively). Among the non-exercise test health status measures, peak VO₂ had

### Table 2. Description, Metric, and Ranges Observed for the Health Status Tools

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Description</th>
<th>Metric (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYHA</td>
<td>Functional classification</td>
<td>4-category scale (1 to 4)</td>
</tr>
<tr>
<td>VSAQ</td>
<td>Pretest estimation of exercise capacity based on symptoms</td>
<td>13-category scale using METS (3 to 10)</td>
</tr>
<tr>
<td>Peak VO₂</td>
<td>Measured oxygen uptake</td>
<td>Continuous variable (6.0–32 mL O₂·kg·min⁻¹)</td>
</tr>
<tr>
<td>Estimated METS</td>
<td>Multiple of RMR estimated from peak speed and grade</td>
<td>Continuous variable (2.4 to 10.4)</td>
</tr>
<tr>
<td>KCCQ</td>
<td>HF-specific health status, measure with 8 domains scale</td>
<td>Transformed 0 to 100</td>
</tr>
<tr>
<td>DASI</td>
<td>Estimate of exercise tolerance</td>
<td>Transformed continuous variable (7 to 38)</td>
</tr>
<tr>
<td>Six-minute walk</td>
<td>Distance walked in 6 minutes</td>
<td>Meters (61 to 507)</td>
</tr>
</tbody>
</table>

NYHA, New York Heart Association functional class; VSAQ, Veterans Specific Activity Questionnaire; RMR, resting metabolic rate; KCCQ, Kansas City Cardiomyopathy Questionnaire; HF, heart failure; DASI, Duke Activity Status Index.
significant associations with the VSAQ score ($r = 0.37$, $P < .05$), but not with the DASI ($r = 0.26$) or NYHA class ($r = -0.14$). Among the KCCQ domains, peak VO$_2$ was significantly associated with quality of life score ($r = 0.46$, $P < .05$), but not with physical limitation, symptom, or clinical summary scores. Six-minute walk distance was significantly though modestly related to peak VO$_2$ ($r = 0.49$, $P < .01$), KCCQ physical limitation ($r = 0.53$, $P < .01$), and clinical summary scores ($r = 0.44$, $P < .05$). Among pretest clinical indices (age, heart rate, EF, FVC, FEV$_1$), only age and EF were related to peak VO$_2$ ($-0.58$, $P < .001$ and 0.46, $P < .01$, respectively). The associations between the various functional and health status measures and peak VO$_2$ were similar when those with nonischemic HF were assessed separately.

### Multivariate Predictors of Peak VO$_2$

Predictors of peak VO$_2$ by multiple regression are presented in Table 6. Age explained the greatest proportion of variance in peak VO$_2$ (49%) followed by the KCCQ quality of life score (23%). Six-minute walk performance was not a significant multivariate determinant of peak VO$_2$. The final regression equation for peak VO$_2$ was

$$\text{peak VO}_2 = 28.8 - 0.33 \times \text{age} + 0.14 \times \text{KCCQ quality of life score}.$$  

The multiple $r$ from the regression model was 0.85 ($P < .001$).

To compare the current sample of HF patients with our previous results among exercise test referrals (largely non-HF) in which peak estimated METS was the dependent variable, a second analysis was performed. Only VSAQ score ($r^2 = 0.34$), age (additional $r^2 = 0.15%$), and KCCQ physical limitation score (additional $r^2 = 16%$) were significant predictors of peak estimated METS (multiple $r$ for model $= 0.81$, $P < .001$) (Table 7).

### Discussion

The present results suggest that commonly-used functional measures, estimates of quality of life, and symptom tools in patients with HF generally have only modest associations with peak VO$_2$ and with one another. The clinical implications of these findings are that: (1) each test or instrument targets a specific aspect of clinical status and they should not be considered interchangeable and (2) none of these measures alone is a reliable surrogate for peak VO$_2$.

Peak VO$_2$ was used as the dependent variable because it is widely recognized as the standard measure of cardiopulmonary function and has wide applications for risk stratification in HF. It should be noted, however, that not all the independent variables employed in the present study were designed specifically to estimate maximal cardiopulmonary function, even though many of them have been used to quantify functional changes in clinical trials. The lack of strong associations between the various functional measures is likely explained in part by the fact that they
assess different aspects of functional status. These observations also highlight the complexity of HF; many factors interact to underlie exercise intolerance in a given patient.1,2,16,20 Symptoms can vary from day to day, and progression of the disease can be unpredictable, leading to a divergence in physiologic function, clinical status, and quality of life.

None of the symptom tools had a strong association with peak VO2; correlation coefficients between peak VO2 and these instruments were in the order of 0.20 to 0.40 (Table 5). The relationship between peak VO2 and the VSAQ (r = 0.37) is similar to what we previously observed using this measure in a more heterogeneous sample of patients referred for exercise testing (r = 0.42).23 In various populations with cardiovascular disease, studies have reported associations between the VSAQ and peak VO2 in the range of 0.46 to 0.57.23–26 In general, the functional questionnaires had stronger associations with estimated than with measured peak VO2. Among the questionnaires, the VSAQ had the strongest association with estimated METS (r = 0.73, P < .001). The latter relationship is similar to what we observed previously among treadmill test referrals, largely without HF (r = 0.79).6

**Association Between Quality of Life and Exercise Tolerance in HF**

Although quality of life would be expected to be closely related to one’s exercise capabilities, studies associating quality of life and exercise tolerance have been mixed. Arena et al21 observed that the Minnesota Living with Heart Failure Questionnaire (MLHF) was significantly related to peak VO2 in patients with HF (r = 0.66, P < .001), though it was poorly related to the VE/VCO2 slope. Koukouvou and colleagues28 reported that the MLHF and other quality of life scores closely paralleled improvements in peak VO2 and other cardiopulmonary exercise test responses before and after 6 months in a rehabilitation program. Conversely, Wilson and colleagues29 reported poor relationships between perceived exercise tolerance as measured by the MLHF and objective measures of circulatory or ventilatory performance. Grigioni et al30 observed poor relationships between the majority of objective clinical markers of disease severity (peak VO2, presence of a third heart sound, signs of fluid retention, electrocardiogram findings, EF, and echocardiographic variables) and the

**Table 6. Explanation of Variance in Peak VO2 by Stepwise Multiple Regression Analysis**

<table>
<thead>
<tr>
<th>Variables Entered</th>
<th>r</th>
<th>r²</th>
<th>New Variance Explained (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.70</td>
<td>0.49</td>
<td>49</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>KCCQ quality of life score</td>
<td>0.65</td>
<td>0.72</td>
<td>23</td>
<td>.001</td>
</tr>
</tbody>
</table>

Regression equation: Peak VO2 = 28.8 – 0.33 (age) + 0.14 (KCCQ quality of life score). KCCQ, Kansas City Cardiomyopathy Questionnaire.

MLHF. Juenger et al31 reported that the physical function domain of the Medical Outcomes Study (SF-36) was significantly related to peak VO2 (r = 0.47, P < .001), but the other 7 SF-36 domains were generally poorly associated with peak VO2 (r = 0.16 to 0.40). In accordance with these studies, peak VO2 had relatively poor relationships with subjective symptom measures in the present study (r = −0.14 to 0.30), although a significant association was observed between peak VO2 and KCCQ quality of life score (r = 0.46, P < .05).

**Association of 6-Minute Walk Performance With Functional Measures**

The 6-minute walk test has been widely used in the evaluation and serial assessment of patients with HF.7,13,32–35 Our HF patients were similar to those in previous studies in that their mean 6-minute walk distance was 328 ± 93 meters.32–35 Previous studies have reported correlation coefficients between 6-minute walk performance and peak VO2 among patients with HF in the range of 0.50 to 0.70.32,34,36–46 The association we observed between 6-minute walk distance and peak VO2 (r = 0.49) was comparable to these previous studies. It might be expected that 6-minute walk performance, as a submaximal measure, would be more closely associated with questionnaires designed to reflect symptom limitations during daily activities than peak VO2. However, although we observed that 6-minute walk distance had a modest association with KCCQ physical limitation score (r = 0.53, P < .01), it was poorly related to KCCQ symptom and quality of life scores. In developing the KCCQ, Green et al8 originally observed a correlation of 0.48 between the physical limitation score and 6-minute walk distance, which is similar to other studies associating various quality of life instruments to 6-minute walk performance.30,31 These studies suggest that although 6-minute walk performance may not be an appropriate surrogate for peak VO2, it is a modest reflection of functional status assessed by questionnaire.

**Table 7. Explanation of Variance in Estimated METS Achieved by Stepwise Multiple Regression Analysis**

<table>
<thead>
<tr>
<th>Variables Entered</th>
<th>r</th>
<th>r²</th>
<th>New Variance Explained (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSAQ</td>
<td>0.58</td>
<td>0.34</td>
<td>34</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Age</td>
<td>0.70</td>
<td>0.49</td>
<td>15</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>KCCQ physical limitation score</td>
<td>0.81</td>
<td>0.65</td>
<td>16</td>
<td>&lt;.05</td>
</tr>
</tbody>
</table>

Regression equation: Estimated METS = 5.1 – 0.07 (age) + 0.37 (VSAQ score) + 0.03 (KCCQ physical limitation score). VSAQ, Veterans Specific Activity Questionnaire; KCCQ, Kansas City Cardiomyopathy Questionnaire.

**Functional Status Measures and EF**

Another potentially useful clinical application relates to whether simple, noninvasive functional measures and exercise test responses are associated with a clinical index such as EF. In previous studies, both 6-minute walk performance
and exercise test responses have been shown to relate poorly to EF and other hemodynamic variables in HF. Although we observed that EF had a stronger relation with peak VO₂ than most previous studies, the relationship was modest (r = 0.46, P < .01). Among the other functional measures, the highest correlation with EF was for 6-minute walk performance, but the relationship was relatively weak (r = 0.28). EF also did not have strong associations with the KCCQ domains, ranging from 0.24 for the physical limitation to 0.54 for the quality of life score. The absence of strong relationships between EF and HF-specific quality of life measures has been observed by other investigators.

Multivariate Analysis

We observed that the multivariate model explained 72% of the variance in peak VO₂, with age accounting for the greatest amount of variance (49%). That age accounted for the greatest amount of variance in peak VO₂ is not surprising, given the decline in peak VO₂ that occurs with age regardless of disease presence, and the greater likelihood of comorbidities and more severe disease in older subjects. Interestingly, the KCCQ quality of life score was the only health status measure that contributed significantly to the model, adding 23% to the variance explained. Similar to our previous study among consecutive exercise test referrals (largely without HF) in which the combination VSAQ score and age were the only significant predictors of estimated exercise capacity (multiple r = 0.79), age and VSAQ score were significant predictors of estimated METs achieved in the present study, along with the KCCQ physical limitation score (multiple r = 0.81). These results suggest that when exercise testing is unavailable, exercise capacity might be reasonably estimated multivariately by considering age and a non-exercise test assessment tool.

Limitations

The study sample size was relatively small. We did not employ practice trials for maximal exercise testing or the 6-minute walk test, both of which are known to improve with repetition. The study was limited to men, and the sample was a heterogeneous group of patients with HF; our findings may not apply to women or specific subgroups of HF. In addition, the sample was a comparatively healthy group of patients with HF, and the results may have differed in patients with more severe HF.

Summary

Health status, including functional capabilities, symptoms associated with daily activities, and physical or psychologic well-being are important indices that are widely used in the assessment of interventions for HF. Accurate and reliable tools that quantify health status are important to appropriately assess responsiveness to therapy in this population. We observed that tools commonly used to assess health and functional status in HF had only modest associations with peak VO₂ and with one-another. In practical terms, this suggests that: (1) laboratory exercise tolerance is not necessarily a good reflection of how patients perceive their capacity to undertake daily activities and (2) cardiopulmonary exercise testing, non-exercise test estimates of physical function, and quality of life measures reflect different facets of health status and one should not be considered a surrogate for another. These findings may have applications for the design and interpretation of clinical trials in which estimates of exercise tolerance are used as efficacy parameters in patients with HF.

References