Peak \( \dot{V}O_2 \)
A Simple yet Enduring Standard

Donna Mancini, MD; Thierry LeJemtel, MD; Keith Aaronson, MD, MS

In this issue of Circulation, Pardaens et al.1 examine the prognostic value of cardiopulmonary exercise testing in ambulatory patients with heart failure who are screened for cardiac transplantation. The investigators failed to demonstrate any significant advantage of ventilatory data over peak oxygen uptake (\( \dot{V}O_2 \)). Over the past 10 years, several investigators have attempted to refine the use of peak \( \dot{V}O_2 \) to improve its prognostic yield. Others have sought methods or equations to estimate peak \( \dot{V}O_2 \) from submaximal data, such as that collected from 6-minute walk tests or during low-level exercise. Most investigators have concluded that the straightforward measurement of peak \( \dot{V}O_2 \) provides the best index of prognosis in patients with ambulatory heart failure.

The measurement of oxygen consumption (\( \dot{V}O_2 \)) in patients with heart failure was first described by Weber et al.2 as a noninvasive method for characterizing cardiac reserve and functional status in these patients. Its use as a prognostic tool has evolved. Szlachcic et al.3 initially described the prognostic use of peak \( \dot{V}O_2 \) in a group of 27 patients. The 1-year mortality rate was 77% for patients with a peak \( \dot{V}O_2 \)<10 mL·kg\(^{-1}\)·min\(^{-1}\) and 21% for those with a peak \( \dot{V}O_2 \) of 10 to 18 mL·kg\(^{-1}\)·min\(^{-1}\).

See p 1152

mL·kg\(^{-1}\)·min\(^{-1}\). Likoff et al.4 reported a 36% mortality rate in 201 patients with heart failure who had a peak \( \dot{V}O_2 \) of 13 mL·kg\(^{-1}\)·min\(^{-1}\); the mortality rate was 15% when peak \( \dot{V}O_2 \) exceeded 13 mL·kg\(^{-1}\)·min\(^{-1}\). Exercise data from the first Veterans Administration Heart Failure Trial (VHeFT)5 also demonstrated that peak \( \dot{V}O_2 \) independently predicted mortality.

To determine whether measuring peak \( \dot{V}O_2 \) could help optimally time cardiac transplantation, we prospectively performed exercise testing with respiratory gas analysis on all ambulatory patients referred to the University of Pennsylvania for cardiac transplantation between October 1986 and December 1989.6 The 116 patients were divided into 3 groups on the basis of their peak \( \dot{V}O_2 \). One group was composed of patients with a peak \( \dot{V}O_2 \)<14 mL·kg\(^{-1}\)·min\(^{-1}\) who were accepted as transplant candidates (n=35). A second group consisted of patients with a peak \( \dot{V}O_2 \)>14 mL·kg\(^{-1}\)·min\(^{-1}\) who had transplant deferred (n=52). Patients with a peak \( \dot{V}O_2 \)<14 mL·kg\(^{-1}\)·min\(^{-1}\) but with a significant comorbidity that precluded transplantation formed a third group (n=27). Age, left ventricular ejection fraction, cause of heart failure, and resting hemodynamic parameters were similar between the groups. One-year survival was 94% in the patient group with a \( \dot{V}O_2 \)>14 mL·kg\(^{-1}\)·min\(^{-1}\). Accepted transplant candidates with a \( \dot{V}O_2 \)<14 mL·kg\(^{-1}\)·min\(^{-1}\) had a 1-year survival of 70%, whereas those patients with a significant comorbidity and reduced \( \dot{V}O_2 \) had a 1-year survival of 47%. Patients accepted for cardiac transplantation had falsely elevated survival because all transplants were treated as a censored observation. If urgent transplant was counted as death, 1-year survival fell to 48%. A peak \( \dot{V}O_2 \)>14 mL·kg\(^{-1}\)·min\(^{-1}\) allowed us to identify patients with severe heart failure whose transplant could be safely deferred. The application of cardiopulmonary stress testing for the selection of potential transplant candidates has subsequently gained widespread acceptance in the United States.7

Peak \( \dot{V}O_2 \) is affected by age, sex, muscle mass, and conditioning status; the percent of predicted peak \( \dot{V}O_2 \) may yield better risk stratification than the absolute value.8 Accordingly, peak \( \dot{V}O_2 \) was measured in 272 patients with advanced congestive heart failure (CHF) who were referred for transplant evaluation.8 Predicted \( \dot{V}O_2 \) was then calculated for each patient using the Astrand and Wasserman equations.8 Patients were then divided into 3 groups on the basis of their peak \( \dot{V}O_2 \); <10, 10 to 14, and >14 mL·kg\(^{-1}\)·min\(^{-1}\). Strata for percent of predicted peak \( \dot{V}O_2 \) were determined by cut points that would yield strata of similar sizes to the above groups. Survival curves for patients stratified by absolute and percent of predicted peak \( \dot{V}O_2 \) were similar. Receiver-operating curves were constructed for absolute peak \( \dot{V}O_2 \) normalized for body weight and percentage of predicted peak \( \dot{V}O_2 \). The area under the curves was roughly equal; therefore, normalization of peak exercise \( \dot{V}O_2 \) for predicted values added minimal prognostic information.

In contrast to our study,8 Stelken et al.9 used multivariate analyses in 181 patients; they found that 50% of predicted peak \( \dot{V}O_2 \) was the most significant predictor of cardiac death (\( P=0.007 \)) and that the area under the curve for percent of predicted peak \( \dot{V}O_2 \) was superior to peak \( \dot{V}O_2 \). However, the prognostic difference between peak \( \dot{V}O_2 \) and percent of predicted peak \( \dot{V}O_2 \) was minimal in both studies. Quite frankly, for middle-aged men with severe heart failure (the stereotypical heart transplant candidate), peak \( \dot{V}O_2 \) is just as prognostically effective as the percent of predicted value. It is only at the extremes of age and, perhaps, sex that the percent of predicted peak \( \dot{V}O_2 \) may provide additional value.
Peak $V_O^2$ is a continuous rather than a discrete variable, and differences in the above 2 studies may be explained by the fact that both groups attempted to assign a threshold or cut-off value to determine transplant candidacy. Stratum-specific likelihood ratios can be used to identify threshold values. Therefore, we calculated stratum-specific likelihood ratios in 140 ambulatory patients referred for cardiac transplant evaluation. The ratios progressively increased as peak $V_O^2$ increased, but no discrete cut point was identified. Therefore, these stratum-specific likelihood ratios indicate that peak $V_O^2$ is a strong and continuous predictor of survival in this population and that it does not have an absolute threshold.

In a further attempt to enhance the predictive power of peak exercise $V_O^2$, investigators have coupled hemodynamic monitoring with oxygen consumption measurements. Chomsky et al evaluated the cardiopulmonary and hemodynamic exercise responses of 185 ambulatory patients with CHF and a mean peak $V_O^2$ of 12.9 mL · kg $^{-1}$ · min $^{-1}$ who were referred for transplant evaluation. They used the following formula to define a normal cardiac output response to exercise: cardiac output = 5 × $V_O^2$ (in L/min) + 3 L/min. On the basis of this formula, these investigators divided their cohort into normal and reduced cardiac output groups. Multivariate analyses found that both a peak $V_O^2$ < 10 mL · kg $^{-1}$ · min $^{-1}$ and a reduced cardiac output response to exercise, as defined by the above equation, were predictive of poor 1-year survival. Whether a straightforward multivariate analysis using directly measured values would have yielded similar findings is unclear. Moreover, the application of invasive hemodynamic monitoring with ventilatory measurements greatly increases the complexity, expense, and potential risks of exercise testing in this population.

We also investigated whether exercise hemodynamic measurements rather than peak $V_O^2$ alone could better identify patients at increased mortality risk. A total of 65 patients underwent bicycle exercise testing with simultaneous metabolic and hemodynamic measurements. Results of multivariate analysis demonstrated that the only exercise variable predictive of survival was left ventricular stroke work index. These results were consistent with those of Griffin et al; they used multiple logistic regression analyses and identified peak exercise stroke work index as the only exercise-derived hemodynamic predictor of mortality. To be accurate, the derivation of peak exercise stroke work requires the absence of mitral regurgitation. Many of our patients with end-stage heart failure have severe mitral regurgitation, which makes the accuracy of this finding questionable.

Most studies fail to demonstrate improved risk stratification with the measurement of exercise hemodynamics beyond the direct measurement of peak $V_O^2$. This is not surprising; the limitations to maximal exercise performance in CHF patients are a mix of both central and peripheral factors. Indeed, the value of peak $V_O^2$ rests in the fact that this measurement integrates elements of cardiac reserve, skeletal muscle, pulmonary, and endothelial dysfunction more than other, traditional prognostic markers in patients with severe heart failure.

Osado et al attempted to further stratify the high-risk group of patients with a peak $V_O^2$ < 14 mL · kg $^{-1}$ · min $^{-1}$ by performing multivariate analyses using all noninvasive exercise parameters measured during exercise testing. Cardiopulmonary exercise testing was performed in 500 patients with CHF who were referred for heart transplantation; 154 patients (31%) had a peak exercise $V_O^2$ ≤ 14 mL · kg $^{-1}$ · min $^{-1}$. Multivariate analyses of exercise and cardiopulmonary variables (ie, peak exercise heart rate, systolic blood pressure, respiratory exchange ratio, minute ventilation, peak $V_O^2$, percent predicted peak $V_O^2$, and anaerobic threshold) were performed to identify the 3-year prognostic risk. Peak systolic blood pressure < 120 mm Hg ($P = 0.0005$) and percent predicted peak $V_O^2$ ≤ 50% ($P = 0.04$) were significant prognostic variables in patients with a peak $V_O^2$ ≤ 14 mL · kg $^{-1}$ · min $^{-1}$.

Survival was 55% at 3 years for the patients with a peak exercise $V_O^2$ ≤ 14 mL · kg $^{-1}$ · min $^{-1}$ and a peak exercise systolic blood pressure < 120 mm Hg; it was 83% in the patients with a systolic blood pressure ≥ 120 mm Hg ($P = 0.004$).

Other investigators, such as Chua et al, examined the clinical and prognostic significance of the ventilatory response to exercise in patients with stable, severe, chronic heart failure. During exercise, a close linear relationship is observed between carbon dioxide production and minute ventilation until the ventilatory threshold is reached, then the slope of this relationship increases. Patients with a steeper ventilatory response have reduced cardiac output during exercise, increased pulmonary pressures, an increased dead space/tidal volume ratio, and potentially augmented chemoreceptor sensitivity. In their analysis of 173 patients, Chua et al found that the ventilatory response to exercise did seem to add prognostic information above that provided by peak $V_O^2$ alone. This study is in contrast to the current study of Pardaens et al. The major differences in these studies illustrate the problem with analyzing all the ancillary data collected during $V_O^2$ testing: the interpretation depends on how you choose to analyze the data. Chua et al used all the data from the start to the end of exercise to derive the ventilatory equivalent for $CO_2$, whereas Pardaens et al used the first 6 minutes of exercise. It does make teleologic sense that the peak $V_O^2$, which more fully reflects the pathophysiologic limitations seen in heart failure, should be a superior prognostic factor than the more focused ventilatory data. Use of submaximal exercise data may be helpful, however, in those patients unable to perform maximal exercise.

One can argue that peak $V_O^2$ is somewhat dependent on patient motivation as well as investigator analysis, but this is less true if the patient reaches a respiratory quotient above 1 and/or the ventilatory threshold. As in most life circumstances, the simpler the approach, the more direct the measurement, and the less manipulation, the better.

Although peak $V_O^2$ is an excellent isolated predictor of outcome, its value can be enhanced by combining it with other important and easily obtainable clinical characteristics. Although peak $V_O^2$ is a useful predictor of prognosis, it should be viewed in the context of the whole clinical presentation. Pretransplant risk stratification was improved by developing a predictive model that incorporated multiple independent predictors of mortality. We developed a Heart Failure Survival Score (HFSS) from 467 ambulatory patients.
with severe CHF who were followed at 2 institutions from July 1986 to September 1994. The model was developed using 268 patients from the University of Pennsylvania hospital who were followed from July 1986 to January 1993, and it was validated in a group of 199 patients at Columbia Presbyterian Hospital who were followed from July 1993 to October 1995.

In this model, 80 clinical variables for each patient that were derived from clinical history, physical examination, and laboratory, exercise, and catheterization data were entered into the data set. Univariate survival analyses were performed using Kaplan Meier analyses. Significant univariate factors were then analyzed with multivariate techniques. Variables were grouped, and those prognostic factors thought to represent different aspects of CHF were incorporated into the model. In the construction of the model, we used clinical judgment to guide the selection process. We specifically sought to include variables that reflect multiple aspects of the pathophysiology of heart failure and that rely minimally on investigator or patient interpretation. The model with the smallest number of variables that could most accurately predict survival was derived.

One statistical model, HFSS, only incorporated noninvasive parameters, including the following 7 variables and their pathophysiological constructs: presence or absence of coronary artery disease (myocardial ischemia), resting heart rate (activation of sympathetic nervous system), left ventricular ejection fraction (the degree of systolic dysfunction), mean arterial blood pressure, presence or absence of intraventricular conduction defects, scoring is based on their pathophysiological construct on baseline ECG (the extent of myocardial fibrosis), serum sodium (the degree of activation of the renin angiotensin system), and peak VO₂. To calculate a prognostic score, the value of the variable and the β-coefficient from the Cox model are multiplied, the products are added, and the absolute value is taken as the HFSS. For noncontinuous variables (ie, coronary artery disease or intraventricular conduction defects), scoring is based on their presence or absence; presence is assigned a value of 1, and absence a value of 0.

Model discrimination was excellent. Stratum-specific likelihood ratios revealed 3 distinct groups in the derivation data set. Patients with prognostic scores >8.1 have excellent survival and do not require transplant listing. Medium-risk (HFSS=7.2 to 8.1) and high-risk (HFSS<7.2) patients have a sufficient mortality risk to warrant transplant listing. Similar survival curves could be generated in the validation sample using the same cut points. Thus, the application of this statistical model, which incorporates several prognostic factors, can help risk-stratify patients more effectively.

Is peak VO₂ by itself as effective as HFSS in predicting outcome? In the model-derivation sample, peak VO₂ did perform as well as the model; however, this was not true in the validation sample.

Is the HFSS the best model for predicting survival in patients with heart failure? Presently, no other multivariate models have been validated prospectively. Whether this score is applicable in the era of β-blockade has not been firmly established.

In conclusion, peak VO₂ is probably the best single measure of prognosis in ambulatory patients with severe heart failure, but risk stratification can be enhanced by using a model that encompasses a variety of prognostic markers, including peak VO₂.

References


Key Words: exercise • heart failure • prognosis • transplantation
Peak \( \text{V} \text{O}_2 \): A Simple yet Enduring Standard
Donna Mancini, Thierry LeJemtel and Keith Aaronson

Circulation. 2000;101:1080-1082
doi: 10.1161/01.CIR.101.10.1080

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2000 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://circ.ahajournals.org/content/101/10/1080