Introduction and objectives. The prognostic value of a single measurement of ejection fraction and peak oxygen uptake in chronic heart failure has been extensively investigated. The aim of our study was to evaluate the prognostic significance of serial changes in ejection fraction and exercise performance in moderate to severe chronic heart failure.

Methods. 182 patients (156 men, 53 [47-58] years) underwent echocardiography and cardiopulmonary exercise testing at baseline and after 10 [8-12] months. Most patients had idiopathic dilated cardiomyopathy (69%) and all patients presented left ventricular ejection fraction <45%. Median follow-up was 21 [14-34] months; cardiac death and heart transplantation were the endpoints. Hazard ratio (HR, per unit) is presented with its 95% confidence interval (CI).

Results. During follow-up 18 patients (9.9%) died and 14 (7.7%) underwent heart transplantation. Baseline ejection fraction (HR, 0.94, 95% CI, 0.89-0.98; P = .006) and mitral regurgitation (HR, 4.22, 95% CI, 1.63-10.92; P = .003), and delta (second examination-baseline) ejection fraction (HR, 0.93, 95% CI, 0.88-0.98; P = .01) were the only significant variables at univariate analysis. Both ejection fraction and delta ejection fraction remained independently associated with events at multivariate analysis. The prognostic power significantly increased between a model including ejection fraction alone and another one including ejection fraction plus delta ejection fraction.

Conclusions. In clinically stable patients with chronic heart failure, ejection fraction and its changes were independently associated with outcome; on the contrary, serial cardiopulmonary exercise testing did not provide significant prognostic value. Baseline plus changes in ejection fraction showed better prognostic performance than baseline ejection fraction alone.

Key words: Heart failure. Prognosis. Echocardiography.
ABBRévIATIONS
CHF: chronic heart failure.
EF: ejection fraction.
VO$_2$: peak oxygen uptake.

INTRODUCTION
Despite the advances in medical treatment and the compelling evidence of multi-drug efficacy in chronic heart failure (CHF), prognosis remains a major concern. The mortality rates remain high and the clinical course is often unpredictable; thus, the detection of prognostic variables is of great importance in the management of CHF patients. The clinical evaluation requires an assessment of the patient’s risk after optimization of medical therapy, in order to provide information about the outcomes and to determine the appropriate allocation of a limited resource such as cardiac transplantation.

The prognostic value of a single measurement of left ventricular ejection fraction (EF) and peak oxygen uptake (VO$_2$) in patients with CHF has been extensively investigated$^{1,2}$; however there are conflicting data on the prognostic significance of their serial measurements. Furthermore it is unclear whether changes of these parameters strictly reflect disease progression and have any impact on late prognosis or symptoms.$^{3,4,12}$

The aim of our study was to evaluate the prognostic significance of serial changes in left ventricular function and exercise performance in patients with moderate to severe CHF, who were in stable clinical conditions.

METHODS
Study Patients

Six-hundred and seven patients with moderate to severe CHF followed in our institution between 1996 and 2000 underwent both echocardiography and cardiopulmonary exercise testing. Inclusion criteria were as follows: stable clinical conditions under optimized medical treatment, echocardiographic left ventricular EF<45%, echocardiogram performed within 2 days from the cardiopulmonary exercise test and repeat evaluation of both tests after a period of ≥4 months (10 (8-12)). According to these criteria, 182 patients were retrospectively selected (156 men, 53 [47-58] years). The majority of patients was in New York Heart Association (NYHA) functional class III (n=149, 82%), and the remaining in class IIb (n=33, 18%). At the first examination medical treatment included digitalis (75%), diuretics (100%), angiotensin converting enzyme inhibitors (96%), amiodarone (33%), and beta-blockers (22%). At the second examination, beta-blockers were added to the standard therapy in a further 55 (31%), thus rising to 53% of the patients population. The etiology of heart failure was idiopathic dilated cardiomyopathy in 125 patients (69%), coronary artery disease in 45 (25%), and valvular disease 12 (6%). Among patients with coronary artery disease, 31 (69%) and 7 (16%) had previous myocardial infarction and previous revascularization by surgery respectively.

Cardiopulmonary Exercise Testing

The exercise tests were performed using an electrically braked bicycle ergometer using a continuous ramp protocol, in which work rate was increased by 10 W/min.

Before each test, oxygen and carbon dioxide analyzers and a flow mass sensor were calibrated by use of available precision gas mixtures and a 3-L syringe, respectively. To stabilize gas measurements, patients were asked to remain still on the ergometer for at least 3 minutes before commencing exercise. A 12-lead electrocardiogram was monitored continuously during the test (Case 16, Marquette Electronics, Milwaukee, Wisconsin, USA), and blood pressure was recorded every 2 minutes by a cuff-manometer. Respiratory gas exchange measurements were obtained breath-by-breath with use of a computerized metabolic cart (Vmax29, Sensormedics, Yorba Linda, California, USA). VO$_2$, carbon dioxide production, minute ventilation, and respiratory exchange ratio were calculated on-line. Peak VO$_2$ was defined as the highest VO$_2$ achieved during exercise and was expressed in milliliters per kilogram per minute (mL/kg/min). The ventilatory anaerobic threshold was determined by the V-slope method and confirmed by ventilatory criteria. Predicted peak VO$_2$ was determined by use of a sex-, age-, height-, and weight-adjusted and protocol-specific formula outlined by Wassermann et al.$^{13}$
Echocardiography

Transthoracic echocardiograms were performed from parasternal long and short axis and apical 4- and 2-chamber views with a Acuson 128XP (Acuson Corporation, Mountain View, California, USA). Left ventricular volume was calculated from orthogonal apical views by use of the area-length method, and EF was derived from the standard equation.

Mitral regurgitation was classified according to 4 grades of color flow Doppler (none or trivial; mild; moderate; severe). The exam was performed by personnel who were unaware of cardiopulmonary exercise test results. The inter-observer variability for repeated measurements of the EF was 5%.

Follow-Up

The patients were followed at the Cardiac Transplant and Heart Failure Unit every 6 months, or by telephone interview with the patient, its family or the primary care physician. Cardiac death and heart transplantation were considered as end-points of the study; median follow-up time was 21 (14-34) months.

Statistical Analysis

Continuous variables are expressed as median and inter-quartile range (I-III) or number (percentage). AT indicates anaerobic threshold; EF, left ventricular ejection fraction; NYHA, New York Heart Association; NS, non significant.

### TABLE 1. Clinical Characteristics, Echocardiographic, and Cardiopulmonary Exercise Testing Data at Baseline/Second Examination (Delta = Absolute Difference Between Second and First Examination)

<table>
<thead>
<tr>
<th></th>
<th>All Cases (n=182)</th>
<th>Event (n=32)</th>
<th>No Event (n=150)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>53 (47-58)</td>
<td>54 (47-58)</td>
<td>53 (46-58)</td>
<td>NS</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>156 (86%)</td>
<td>30 (94%)</td>
<td>126 (84%)</td>
<td>NS</td>
</tr>
<tr>
<td>Etiology of heart failure, n (%)</td>
<td>125 (69%)</td>
<td>24 (75%)</td>
<td>101 (67%)</td>
<td>–</td>
</tr>
<tr>
<td>Dilated cardiomyopathy</td>
<td>12 (6%)</td>
<td>1 (3%)</td>
<td>11 (7%)</td>
<td>NS</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>45 (25%)</td>
<td>7 (22%)</td>
<td>38 (25%)</td>
<td>NS</td>
</tr>
<tr>
<td>Valvular disease</td>
<td>12 (6%)</td>
<td>1 (3%)</td>
<td>11 (7%)</td>
<td>NS</td>
</tr>
<tr>
<td>III NYHA class, n (%)</td>
<td>145 (80%)</td>
<td>25 (78%)</td>
<td>120 (80%)</td>
<td>NS</td>
</tr>
<tr>
<td>Beta-blockers, n (%)</td>
<td>41 (22%)</td>
<td>6 (19%)</td>
<td>35 (23%)</td>
<td>NS</td>
</tr>
<tr>
<td>EF, %</td>
<td>29 (23-35)</td>
<td>25 (19-31)</td>
<td>30 (23-35)</td>
<td>.006</td>
</tr>
<tr>
<td>Mitral regurgitation, n (%)</td>
<td>103 (57%)</td>
<td>11 (34%)</td>
<td>92 (61%)</td>
<td>–</td>
</tr>
<tr>
<td>None or trivial</td>
<td>63 (34%)</td>
<td>14 (44%)</td>
<td>49 (33%)</td>
<td>NS</td>
</tr>
<tr>
<td>Mild</td>
<td>16 (9%)</td>
<td>7 (22%)</td>
<td>9 (6%)</td>
<td>.003</td>
</tr>
<tr>
<td>Moderate</td>
<td>161 (86.5-19.3)</td>
<td>16.6 (12.9-18.5)</td>
<td>16.3 (13.9-19.3)</td>
<td>NS</td>
</tr>
<tr>
<td>Peak VO2, mL/kg/min</td>
<td>10.3 (8.8-12.5)</td>
<td>10.1 (8.7-11.8)</td>
<td>10.8 (8.8-12.7)</td>
<td>NS</td>
</tr>
<tr>
<td>Exercise duration, min</td>
<td>9 (7-11)</td>
<td>8 (7-11)</td>
<td>9 (7-11)</td>
<td>NS</td>
</tr>
<tr>
<td>Delta EF, %</td>
<td>1.5 (0.0 to 4.0)</td>
<td>2.0 (0.0 to 6.0)</td>
<td>0.0 (0.0 to 0.0)</td>
<td>.01</td>
</tr>
<tr>
<td>Delta mitral regurgitation, n (%)</td>
<td>71 (39%)</td>
<td>18 (56%)</td>
<td>53 (35%)</td>
<td>–</td>
</tr>
<tr>
<td>≤1, improved</td>
<td>22 (12%)</td>
<td>3 (10%)</td>
<td>19 (13%)</td>
<td>NS</td>
</tr>
<tr>
<td>&gt;1, worsened</td>
<td>89 (49%)</td>
<td>11 (34%)</td>
<td>78 (52%)</td>
<td>NS</td>
</tr>
<tr>
<td>Delta peak VO2, mL/kg/min</td>
<td>0.3 (0.2 to 2.5)</td>
<td>0.3 (0.2 to 2.5)</td>
<td>0.3 (0.2 to 2.5)</td>
<td>NS</td>
</tr>
<tr>
<td>Delta AT, mL/kg/min</td>
<td>0.3 (0.0 to 1.7)</td>
<td>0.1 (0.0 to 1.5)</td>
<td>0.1 (0.0 to 1.5)</td>
<td>NS</td>
</tr>
<tr>
<td>Delta exercise duration, min</td>
<td>0.0 (0.0 to 2.0)</td>
<td>0.0 (0.0 to 2.0)</td>
<td>0.0 (0.0 to 2.0)</td>
<td>NS</td>
</tr>
<tr>
<td>Beta-blockers 2nd examination, n (%)</td>
<td>96 (53%)</td>
<td>16 (30%)</td>
<td>80 (53%)</td>
<td>NS</td>
</tr>
</tbody>
</table>

Data are expressed as median and inter-quartile range (I-III) or number (percentage). AT indicates anaerobic threshold; EF, left ventricular ejection fraction; NYHA, New York Heart Association; NS, non significant.
Statistical significance was settled at a P value <.05. The statistical analyses were carried out with the Statistical Package for the Social Sciences release 10.0 for Windows and with the S-PLUS (S-PLUS 2000 Guide to Statistics, Volume 2, Data Analysis Products Division, MathSoft, Seattle, WA).

RESULTS
During the follow-up period of 21 (14-34) months, 18 patients (9.9%) died and 14 (7.7%) underwent heart transplantation. Causes of cardiac death were refractory heart failure (60%) and arrhythmic death. The main echocardiography and cardiopulmonary exercise testing results in the overall population, in patients with and without events are listed in Table 1. EF (HR, 0.94; 95% CI, 0.89-0.98; P=.006) and mitral regurgitation (HR, 4.22; 95% CI, 1.63-10.92; P=.003) at first examination, and delta EF (HR, 0.93; 95% CI, 0.88-0.98; P=.01) were the only significant variables associated with events at univariate analysis. EF and delta EF, even adjusted for the main prognostic factors of CHF cited in literature, remained independently associated with events at multivariate analysis (Table 2). The models including EF and another one including EF plus delta EF showed different \( \chi^2 \) value (7.77 and 19.09, respectively), suggesting an incremental statistical power in outcome prediction with the echocardiographic follow-up. Similarly, model validation with bootstrap technique showed a better performance of D, U, Q indexes of the EF plus delta EF model with respect to the model with EF alone (Table 2).

DISCUSSION
The major findings of this study are that in clinically stable patients with moderate to severe CHF: a) baseline EF is an independent predictor of events, and b) changes in EF are independently associated with outcome as well, and provide incremental prognostic power over EF alone.

Echocardiographic Findings
A single measurement of EF has already been shown as a strong predictor of prognosis in CHF.1-4 Similarly, in our study a low baseline left ventricular EF was independently associated with events. Less is known about the prognostic significance of the changes over time in left systolic function. We observed that changes in EF are independently associated with outcome and provide an adjunctive prognostic power over EF alone. This finding was confirmed by Cintorion et al from the V-HeFT studies in which sequential changes in left ventricular EF provided additional prognostic information, even adjusted for baseline EF and medical therapy. On the contrary, Gullestad et al10 did not observe any differences in outcome between patients whose EF increased when compared to those whose decreased during follow-up; it should be noted, however, that repeated left ventricular EF evaluations were not obtained in all patients and that different imaging methods were used, with value differences between the different techniques. A lack of relationship between changes in EF and events was also found by Florea et al in a small sample size12; the patients’ selection criteria (mild to moderate CHF) or the different methods in calculating volume and ejection fraction (from M-mode dimensions using cubed formula versus area-length method in our study) may explain the discrepancies observed.

Cardiopulmonary Exercise Findings
In our study baseline and changes in exercise tolerance during serial examinations did not distinguish patients with from patients without events. These findings are in contrast with some authors; in fact, serial changes in peak VO2 distinguished patients at high risk of events from those with sufficiently stable clinical conditions, that can be removed from the transplant list. Stevenson et al10 reported that out of 107 patients, 31 in stable clinical status who increased their peak VO2 by at least 2 mL/min/kg were removed from the waiting list. The short-term survival of this group was not significantly different from patients who had received heart transplantation. Also Levine et al15 observed that patients with the most favourable outcome had, over approximately 2 years, a significant increase in peak VO2 from 12.2±1.3 to 18.7±5.3 mL/min/kg. Florea et al12 similarly found that patients with increased peak VO2 over time showed a better prognosis at 2 years than those with a decrease in peak VO2.

Patient selection, baseline peak VO2 values, severity of CHF, and drug regimen may explain the discrepancy between other and our results. The studies of Levine et al15 and Stevenson et al14 included patients with end-stage heart failure already on the list for transplantation, with peak VO2<14 mL/min/kg; the study by Florea et al12 dealt with a patient population with mild to moderate CHF with mean peak VO2 of 18±6 mL/min/kg. On the contrary, we included patients with moderate to severe CHF who underwent serial examination with a mean peak VO2 of 16.7±4.3 mL/min/kg (intermediate exercise capacity).

The effect of beta-blockers on maximal exercise tolerance (peak VO2) in patients with CHF has been investigated,16-18 and it has recently been suggested that its prognostic value under beta-blocker treatment should be re-evaluated.19 In agreement with our results, Gullestad et al20 demonstrated that in clinically
stable patients with moderate to severe heart failure receiving beta-blocker treatment, changes in peak VO\(_2\) did not yield additional predictive information. The recommended medical treatment for this patient group evolved during the study period, thus the use of beta-blockers increased from 22% at the first examination to 53% at the second examination. Nevertheless, treatment with beta-blockers showed no relation with EF, delta EF, peak VO\(_2\), delta peak VO\(_2\) (data not shown), and with outcome prediction, since multivariate analysis was also adjusted for this variable.

**Limitations**

This is a retrospective study based on a small sample with respect to the overall population followed in our Institution. Inclusion criteria, i.e. first and second examination including 2 tests within 2 days of each other and the exclusion of patients who died or required transplantation before the second examination, leads to a reduction in the sample size. Furthermore, baseline exercise testing involved only 1 test; an improved exercise capacity could result in part from familiarization with the procedure. Our patients, in stable clinical conditions, were predominantly on NYHA class III (82%). This probably accounted for the lack of relationship between NYHA class and events, in contrast with other studies. Data on left ventricular diastolic function were not available for all patients.

**CLINICAL IMPLICATIONS**

Our findings, obtained from a head to head comparison of echocardiography and cardiopulmonary exercise testing in the same patients, may have implications for the every-day clinical management of CHF. Changes in EF are an easily measurable and clinically important variable in a subset of patients with intermediate functional capacity, whilst the prognostic power of cardiopulmonary exercise testing, a technical demanding and expensive test, seems to be limited. Thus, a rational and pragmatic risk stratification should include EF and, in particular, the measurement of its changes, since it yields an adjunctive predictive power of late outcome.
REFERENCES

8. van den Breuck SAJ, van Veldhuisen DJ, de Graff PA, Landsman MLJ, Lie KI. Comparison between New York Heart Association Classification and peak oxygen consumption in the assessment of functional status and prognosis in patients with mild to moderate chronic congestive heart failure secondary to either ischemic or idiopathic dilated cardiomyopathy. Am J Cardiol. 1992;70:359-63.