

# Cardiopulmonary Exercise Testing Provides Additional Prognostic Information in Cystic Fibrosis

Helge Hebestreit<sup>1</sup>, Erik H. J. Hulzebos<sup>2</sup>, Jane E. Schneiderman<sup>3</sup>, Chantal Karila<sup>4</sup>, Steven R. Boas<sup>5</sup>, Susi Kriemler<sup>6</sup>, Tiffany Dwyer<sup>7,8,9</sup>, Margareta Sahlberg<sup>10</sup>, Don S. Urquhart<sup>11</sup>, Larry C. Lands<sup>12</sup>, Felix Ratjen<sup>3</sup>, Tim Takken<sup>2</sup>, Liobou Varanistkaya<sup>1</sup>, Viktoria Rücker<sup>13</sup>, Alexandra Hebestreit<sup>1</sup>, Jakob Usemann<sup>14</sup>, and Thomas Radtke<sup>6</sup>; for the Prognostic Value of CPET in CF Study Group

<sup>1</sup>University Children's Hospital, Wuerzburg, Germany; <sup>2</sup>Child Development & Exercise Center, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, the Netherlands; <sup>3</sup>Division of Respiratory Medicine, Department of Pediatrics, Hospital for Sick Children, Toronto, Ontario, Canada; <sup>4</sup>Service de Pneumologie et Allergologie Pédiatriques, Centre de Ressources et Compétences dans la Mucoviscidose, Hôpital Necker Enfants Malades, Université Paris V – Descartes, Paris, France; <sup>5</sup>Northwestern University Feinberg School of Medicine, Chicago, Illinois; <sup>6</sup>Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland; <sup>7</sup>Discipline of Physiotherapy, Faculty of Health Sciences and <sup>9</sup>Central Clinical School, Sydney Medical School, University of Sydney, Sydney, Australia; <sup>8</sup>Department of Respiratory and Sleep Medicine, Royal Prince Alfred Hospital, Sydney, Australia; <sup>10</sup>Department of Pediatrics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Goteborg, Sweden; <sup>11</sup>Department of Paediatric Respiratory and Sleep Medicine, Royal Hospital for Sick Children, Edinburgh, United Kingdom; <sup>12</sup>Montreal Children's Hospital, McGill University Health Centre, Montreal, Quebec, Canada; <sup>13</sup>Institute for Clinical Epidemiology and Biometry, Wuerzburg, Germany; and <sup>14</sup>University Children's Hospital Basel, Basel, Switzerland

ORCID ID: 0000-0002-1723-1070 (T.R.).

## Abstract

**Rationale:** The prognostic value of cardiopulmonary exercise testing (CPET) for survival in cystic fibrosis (CF) in the context of current clinical management, when controlling for other known prognostic factors, is unclear.

**Objectives:** To determine the prognostic value of CPET-derived measures beyond peak oxygen uptake ( $\dot{V}O_{2peak}$ ) following rigorous adjustment for other predictors.

**Methods:** Data from 10 CF centers in Australia, Europe, and North America were collected retrospectively. A total of 510 patients completed a cycle CPET between January 2000 and December 2007, of which 433 fulfilled the criteria for a maximal effort. Time to death/lung transplantation was analyzed using Cox proportional hazards regression. In addition, phenotyping using hierarchical Ward clustering was performed to characterize high-risk subgroups.

**Measurements and Main Results:** Cox regression showed, even after adjustment for sex, FEV<sub>1</sub>% predicted, body mass index (z-score), age at CPET, *Pseudomonas aeruginosa* status, and CF-related diabetes as covariates in the model, that  $\dot{V}O_{2peak}$  in % predicted (hazard ratio [HR], 0.964; 95% confidence interval [CI], 0.944–0.986), peak work rate (% predicted; HR, 0.969; 95% CI, 0.951–0.988), ventilatory equivalent for oxygen (HR, 1.085; 95% CI, 1.041–1.132), and carbon dioxide (HR, 1.060; 95% CI, 1.007–1.115) (all  $P < 0.05$ ) were significant predictors of death or lung transplantation at 10-year follow-up. Phenotyping revealed that CPET-derived measures were important for clustering. We identified a high-risk cluster characterized by poor lung function, nutritional status, and exercise capacity.

**Conclusions:** CPET provides additional prognostic information to established predictors of death/lung transplantation in CF. High-risk patients may especially benefit from regular monitoring of exercise capacity and exercise counseling.

**Keywords:** cystic fibrosis; prognosis; peak oxygen uptake; exercise testing; lung transplantation

(Received in original form June 16, 2018; accepted in final form October 12, 2018)

A complete list of the Prognostic Value of CPET in CF Study Group members may be found before the beginning of the REFERENCES.

Author Contributions: Conception and design, H.H. Acquisition of data, A.H., D.S.U., C.K., E.H.J.H., H.H., J.E.S., L.C.L., L.V., M.S., S.K., S.R.B., T.T., and T.D. Genotype classification, T.R. and H.H. Statistical analysis, J.U., T.R., and V.R. Interpretation, F.R., H.H., J.U., and T.R. First draft, H.H. and T.R. All authors edited, reviewed, and approved the final version of the manuscript.

Correspondence and requests for reprints should be addressed to Helge Hebestreit, M.D., Universitäts-Kinderklinik, Josef-Schneider-Strasse 2, 97080 Würzburg, Germany. E-mail: hebestreit@uni-wuerzburg.de.

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org).

Am J Respir Crit Care Med Vol 199, Iss 8, pp 987–995, Apr 15, 2019

Copyright © 2019 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201806-1110OC on October 15, 2018

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

## At a Glance Commentary

### Scientific Knowledge on the

**Subject:** Cardiopulmonary exercise testing is a predictor of survival in cystic fibrosis, but the evidence is based on studies with small sample sizes that could not adjust for important confounders. Moreover, an extensive evaluation of the prognostic utility of various cardiopulmonary exercise testing parameters and their integration into cluster analysis has not been previously undertaken.

### What This Study Adds to the

**Field:** This large, international multicenter study extends previous knowledge on the prognostic role of traditional cardiopulmonary exercise testing parameters and identifies the ventilatory equivalent for oxygen and carbon dioxide at peak exercise as important predictors of the compound outcome survival/lung transplantation. Using an unbiased, data-driven clustering approach, we identified a high-risk phenotype with poor lung function, nutritional status, and substantially reduced exercise capacity: a subgroup of patients who may especially benefit from regular monitoring of exercise capacity and exercise counselling.

In patients with cystic fibrosis (CF), an increased risk of death has been associated with poor values of FEV<sub>1</sub>; short stature; low body mass index (BMI); pancreatic insufficiency; CF-related diabetes (CFRD); and chronic infection with specific pathogens, such as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, or *Mycobacteria* other than tuberculosis (1–3).

Several studies have shown that cardiopulmonary exercise testing (CPET) can also provide prognostic information in CF with respect to mortality (4–9).  $\dot{V}_{O_2\text{peak}}$  and change in  $\dot{V}_{O_2\text{peak}}$  over time, peak work rate ( $W_{\text{peak}}$ ), the respiratory equivalent for oxygen ( $\dot{V}_E$  divided by  $\dot{V}_{O_2}$ ) at peak exercise ( $\dot{V}_E/\dot{V}_{O_2\text{peak}}$ ), and the breathing reserve index at the ventilatory threshold ( $\dot{V}_E$  divided by the estimated maximal voluntary ventilation [MVV]) during an incremental cycling test were identified as predictors of death (4–9).

The analysis of prognostic factors in CF remains a challenge because of its multifactorial complexity. Previous studies investigating the association between CPET and survival in CF usually considered only a selection of established predictors of mortality and/or lung transplantation (LTx) (e.g., FEV<sub>1</sub> and BMI) for adjustment in their statistical models. The decision for the choice of covariates for the final multivariable models was mostly based on significant univariate associations of single CPET parameters. This may, however, overestimate the predictive value of CPET for CF survival because numerous other factors are associated with the outcome.

Another possibility of identifying high-risk patients is to use cluster analysis, such as Ward hierarchical clustering, an unbiased, data-driven approach to define clinical/physiologic phenotypes (10). Such an analysis may help to characterize subgroups with respect to clinical characteristics that could be useful for further investigations and/or intervention strategies.

The primary objective of the present project was to determine the prognostic value of  $\dot{V}_{O_2\text{peak}}$  (primary analysis) and other CPET-derived parameters after rigorous adjustment of a significant number of established predictors of mortality in CF that are usually assessed during routine clinical assessments in a large multicenter cohort. We further wanted to evaluate whether the importance of predictors changes over time using prespecified time periods (i.e., 5, 8, and 10 yr after CPET) to allow comparisons with previous studies (4–7). An additional objective in relation to the study question was to examine whether CPET-derived variables were important to define phenotypes that are associated with LTx and mortality.

## Methods

### Study Design and Subjects

For this retrospective study, we analyzed data of patients with CF aged greater than or equal to 10 years who had a full CPET meeting prespecified criteria between January 1, 2000, and December 31, 2007, and for whom follow-up information on survival or LTx was available 5 years after CPET. Patients who left their respective CF center earlier than 5 years after CPET without information on LTx or subsequent survival available were not included in the analysis. Ethical approval was obtained from

all respective ethical research committees, if required (see online supplement).

### Cardiopulmonary Exercise Testing

For each patient, only one CPET was included in the analysis. In patients who had multiple tests during the study period, the first valid test was selected. An overview on equipment and exercise protocols used to perform the CPET (see Table E1 in the online supplement) and criteria for a maximal effort are available in the online supplement.

### Statistical Analysis

Data are presented as number (%), mean  $\pm$  SD (ranges), and hazard ratios (HRs) with 95% confidence intervals (CI). Data from lung function testing and CPET were converted to % predicted (11–13). We calculated BMI z-scores for children and adolescents (14) and adults (15). The primary composite endpoint for survival was LTx and/or death.

The relationship of  $\dot{V}_{O_2\text{peak}}$  (and  $W_{\text{peak}}$ ) with LTx and/or death was visualized by plotting survival (Kaplan-Meier curves) of three groups using cutoffs used by Nixon and colleagues (4) for Cox proportional hazards regression. The Cox proportional hazards assumption was verified graphically (log-log plot) and tested using Schoenfeld residuals and there was no indication of a violation of this assumption.

First, the association between potential predictors and the compound outcome (death and/or LTx) was studied in a univariate analysis using Cox proportional random effects hazards models adjusted for clustering on the center level. In a second step, we assessed the prognostic value of CPET-derived parameters ( $\dot{V}_{O_2\text{peak}}$ ,  $W_{\text{peak}}$ ,  $\dot{V}_E/\text{MVV}_{\text{pred}}$ ,  $\dot{V}_E/\dot{V}_{O_2\text{peak}}$ ,  $\dot{V}_E/\dot{V}_{CO_2\text{peak}}$ , oxygen saturation at peak exercise as measured by pulse oximetry ( $Sp_{O_2\text{peak}}$ ), and the  $\Delta\dot{V}_E/\Delta\dot{V}_{CO_2}$  slope) in addition to known predictors of survival in CF in multivariable models. Cox proportional hazards regression models were performed including FEV<sub>1</sub> (% predicted), BMI (z-score), age at CPET, sex, and the binary coded variables (yes/no) chronic *P. aeruginosa* colonization and CFRD. Exocrine pancreatic insufficiency was a predefined covariate to be considered for the statistical model. However, because none of the cases (LTx/death) in our dataset was pancreatic sufficient, we were not able to compute effect estimates and 95% CIs. Consequently, pancreatic insufficiency was not

included in the statistical models. Because data on cystic fibrosis transmembrane conductance regulator (CFTR) genotype were missing for 22% of patients with unequal distribution of unavailable/unidentified mutations among cases and survivors (52.7% vs. 5.3%) (see Table E2) and there was no significant difference with respect to the distribution of mutations between cases and control subjects for patients in whom the genotype was available, we did not consider CFTR genotype as a covariate in the final models.

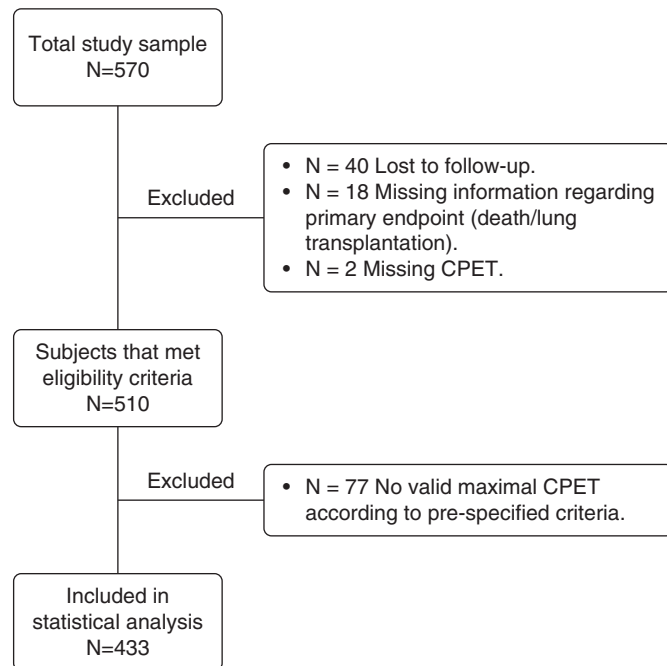
The final Cox model analysis was restricted to 10-year follow-up data because of low numbers of cases ( $n = 8$ ) from only three study centers with more than 10-year follow-up data. In addition, we evaluated predictors of survival for different time periods (e.g., 5, 8, and 10 yr) using Cox models with time-varying covariates.

In addition, an unsupervised, data-driven approach was used to explore the relevance of CPET-derived parameters to predict death/LTx in patients with CF. Physiologic and clinical parameters were used as input to define clusters by Ward hierarchical clustering (10). The relevance of each clustering parameter was investigated using a Forest plot. Based on the results of the Forest plot, the parameters for clustering were selected and finally, the association of the identified clusters with LTx and/or death using Kaplan-Meier curves was studied.

All analyses were performed using Stata statistical software version 12 (StataCorp) and R version 3.0.2 (<http://www.R-project.org>). A probability for a type I error  $P$  less than 0.05 was considered statistically significant.

## Results

Data from 10 CF centers from Australia, Europe, and North America were included in this study. A flow chart is shown in Figure 1. A total of 433 of 510 CPETs fulfilled at least one of the four predefined criteria for a maximal effort (i.e., plateauing of  $\dot{V}_{O_2}$ , or respiratory exchange ratio, heart rate, or  $\dot{V}_E$  exceeding the previously mentioned thresholds) and were, thus, used for the final analysis. Patients with a nonmaximal effort during CPET ( $n = 77$ ) were younger ( $13.7 \pm 3.5$  yr;  $P < 0.001$ ), had comparable FEV<sub>1</sub> ( $74.0 \pm 21.7\%$  predicted;  $P = 0.868$ ), and showed a trend of lower  $\dot{V}_{O_2\text{peak}}$  ( $77.3 \pm 18.6\%$  predicted;  $P = 0.066$ ) compared with the



**Figure 1.** Flow chart of included patients. CPET = cardiopulmonary exercise test.

final study population. Of 77 excluded patients, eight (10%) died or had an LTx during the entire follow-up period. We noticed differences in deaths/LTx among study centers (see Table E3) and consequently adjusted all models for study center. Table 1 summarizes patients' characteristics. Detailed information on patient's clinical characteristics and CPET outcomes separated by study center are shown in Tables E3 and E4.

Survival rates in our study population were 93.1%, 84.7%, and 69.3% after 5, 8, and 10 years, respectively. Survival probabilities for different  $\dot{V}_{O_2\text{peak}}$  categories are shown in the Kaplan-Meier curve (Figure 2). A figure with Wpeak categories is shown in the online supplement (see Figure E1).

### Results of Cox Proportional Hazards Models (Univariate and Adjusted Analyses)

In univariate random effects models with study center adjustment, the CPET-derived information on  $\dot{V}_{O_2\text{peak}}$ , Wpeak,  $\dot{V}_E/\dot{V}_{O_2\text{peak}}$ ,  $\dot{V}_E/\dot{V}_{CO_2\text{peak}}$ ,  $\dot{V}_{E\text{peak}}/MVV_{\text{pred}}$ , Sp<sub>O<sub>2</sub></sub>peak, and the  $\Delta\dot{V}_E/\Delta\dot{V}_{CO_2}$  slope was each a significant predictor of death or LTx (see Table E5). Furthermore, in the complete data set, all CPET-derived parameters were significant in the univariable models without study center

adjustment (crude estimates are shown in Table E6). A low FEV<sub>1</sub> and BMI, as well as chronic *P. aeruginosa* colonization, were also related to death or LTx (see Tables E5 and E6).

A subanalysis restricted to patients with advanced lung disease (FEV<sub>1</sub>  $\leq 40\%$  predicted) identified age at CPET,  $\dot{V}_{O_2\text{peak}}$ , Wpeak,  $\dot{V}_E/\dot{V}_{O_2\text{peak}}$ , and  $\dot{V}_E/\dot{V}_{CO_2\text{peak}}$  as significant predictors of death or LTx (see Table E7). Comparable results for CPET variables were found in non-center-adjusted analyses (see Table E8).

In an exploratory analysis aiming to assess the short-term prognostic value of CPET-derived parameters and restricted to 2-year follow-up data ( $n = 433$  patients with 11 death/LTx cases),  $\dot{V}_{O_2\text{peak}}$ , Wpeak,  $\dot{V}_E/\dot{V}_{O_2\text{peak}}$ ,  $\dot{V}_E/\dot{V}_{CO_2\text{peak}}$ ,  $\dot{V}_{E\text{peak}}/MVV_{\text{pred}}$ , and Sp<sub>O<sub>2</sub></sub>peak as well as FEV<sub>1</sub> and BMI were each significantly associated with death or LTx (see Tables E9 and E10). Results from these short-term models were similar to complete dataset models (see Tables E5 and E6), except for a significant effect of chronic *P. aeruginosa* colonization on death or LTx (see Tables E9 and E10).

When adjusting the models for sex, age at CPET, FEV<sub>1</sub> (% predicted), BMI (z-score), chronic *P. aeruginosa* colonization, and CFRD the variables  $\dot{V}_{O_2\text{peak}}$ , Wpeak,

**Table 1.** Patient Characteristics

Mean follow-up, all patients, yr	8.9 ± 2.9 (0.1 to 14.0)
Mean follow-up, survivors, yr	9.6 ± 2.4 (5.0 to 14.0)
Mean follow-up, death or LTx, yr	5.9 ± 3.3 (0.1 to 13.5)
Death or LTx, n (%)	74 (17.1)
Age, yr	16.6 ± 6.1 (10.0 to 44.5)
Sex, female, n (%)	184 (42.5)
BMI (z-score)	-0.70 ± 1.0 (-4.53 to 1.89)
FEV <sub>1</sub> % predicted	73.4 ± 21.8 (19.7 to 123.4)
Genotype, n (%)	
CFTR, both alleles from classes I-III*	315 (72.7)
CFTR, at least one allele from classes IV-V*	22 (5.1)
CFTR, at least one allele unknown/not available	96 (22.2)
Comorbidities, n (%)	
Chronic <i>Pseudomonas aeruginosa</i> infection	295/424 (69.6)
Pancreatic insufficiency	404/430 (93.3)
CFRD	34/426 (8.0)
Exercise capacity	
$\dot{V}_{O_2}$ peak % predicted	82.0 ± 21.0 (23.3 to 151.2)
Wpeak % predicted	91.2 ± 23.2 (17.0 to 197.0)

*Definition of abbreviations:* BMI = body mass index; CFRD = cystic fibrosis–related diabetes; CFTR = cystic fibrosis transmembrane conductance regulator; LTx = lung transplantation;  $\dot{V}_{O_2}$ peak = peak oxygen uptake; Wpeak = peak work rate.

N = 433. Data are mean ± SD (ranges) or number (% of the study sample).

\*Information on both CFTR mutations were available for 337 patients.

$\dot{V}_E/\dot{V}_{O_2}$ peak, and  $\dot{V}_E/\dot{V}_{CO_2}$ peak remained significant predictors of death or LTx (Table 2). Characteristics of patients within the three  $\dot{V}_{O_2}$ peak categories are shown in Table E11. Patients in the highest and middle  $\dot{V}_{O_2}$ peak groups ( $\geq 82\%$  predicted and 59–81% predicted, respectively) had a 72% (HR, 0.278; 95% CI, 0.088–0.882;  $P = 0.030$ ) and 49% (HR, 0.507; 95% CI, 0.259–0.993;  $P = 0.048$ ) lower risk of dying or to receive a lung transplant within the next 10 years compared with patients with a  $\dot{V}_{O_2}$ peak of less than or equal to 58% of the predicted value. In the model with Wpeak,

patients in the highest Wpeak group ( $\geq 92\%$  predicted) showed a trend for lower chance for death or LTx (HR, 0.417; 95% CI, 0.155–1.123;  $P = 0.084$ ).

In the Cox models including time-varying covariates, no effects were observed for different predictors of survival for different time periods (e.g., 5, 8, and 10 yr, data not shown).

### Results of Ward Hierarchical Clustering

All CPET parameters and FEV<sub>1</sub> had a high variable importance to define clusters, as

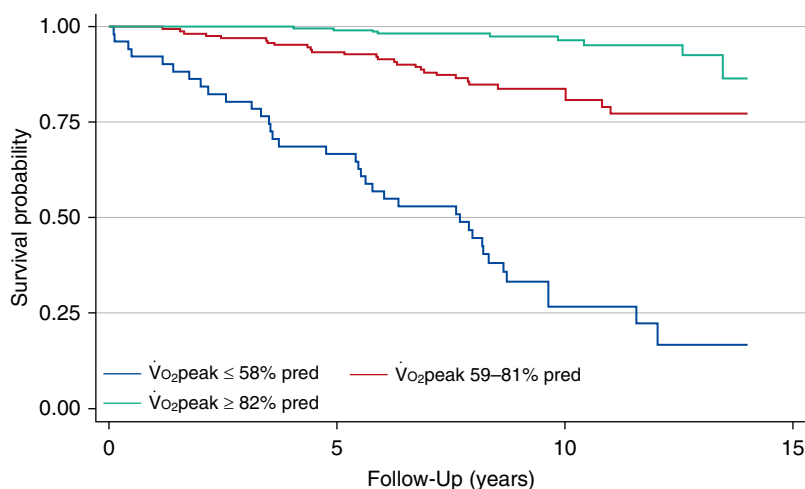
indicated by the Forest plot (Figure 3). However, the binary coded variables (sex, chronic *P. aeruginosa* infection, pancreatic insufficiency, and CFRD) were less important. Based on the forest plot, the eight continuous variables (FEV<sub>1</sub>, BMI, Wpeak, age at CPET,  $\dot{V}_{O_2}$ peak,  $\dot{V}_E/\dot{V}_{O_2}$ peak,  $\dot{V}_E/\dot{V}_{CO_2}$ peak, and  $\dot{V}_E/\dot{V}_{CO_2}$ peak) were introduced into the principal component analysis, and five orthogonal factors explained more than 95% of variance (see Table E12). We identified four clusters (Table 3). The prevalence of the primary outcome death/LTx during the study period was 2% in cluster 1, 15% in cluster 2, 67% in cluster 3, and 39% in cluster 4. Cluster 3 was considered a high-risk cluster because the prevalence of the primary outcome death/LTx during the 10-year study period (63%) was highest compared with the other three clusters (Table 3, Figure 4). This cluster was further characterized by a poor performance for all CPET parameters, a high prevalence of *P. aeruginosa* colonization, low BMI z-score, and low FEV<sub>1</sub> and FVC values (see Figure E2).

We further identified a group of older patients ( $n = 33$ , cluster 4) with poor lung function, high prevalence of chronic *P. aeruginosa* colonization and CFRD, but only a modest reduction in BMI and (almost) preserved exercise capacity (Table 3). This group of patients had better survival compared with patients belonging to cluster 3.

## Discussion

Previous studies of the prognostic value of CPET in CF have been limited by small sample sizes that could not control for many well-recognized confounding prognostic factors (4–9). The present study evaluated a large international sample of patients to examine the utility of CPET parameters in predicting survival over the subsequent 10 years. The study confirms the importance of  $\dot{V}_{O_2}$ peak and Wpeak as key predictors of survival, but also identifies other CPET measures ( $\dot{V}_E/\dot{V}_{O_2}$ peak and  $\dot{V}_E/\dot{V}_{CO_2}$ peak) that may be of great prognostic significance. The use of cluster analysis further suggests at-risk phenotypes in whom early recognition, nutritional counseling, and exercise intervention could be most beneficial.

This is the largest study investigating the prognostic value of CPET-derived



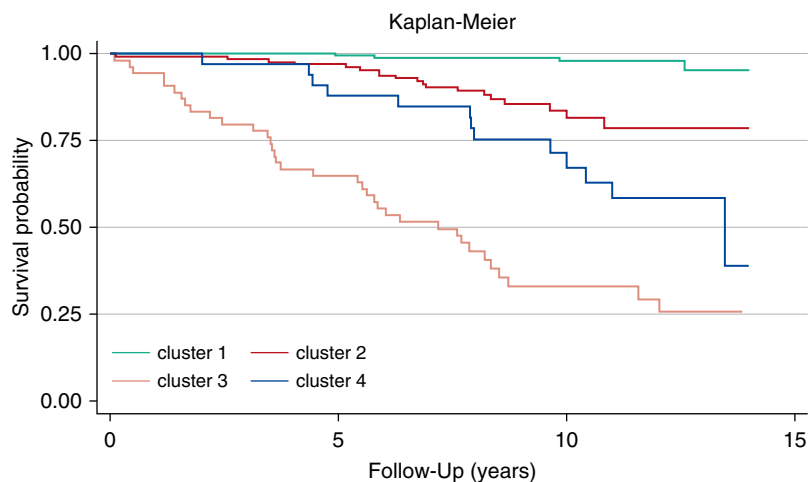
**Figure 2.** Kaplan-Meier survival curve for three different  $\dot{V}_{O_2}$ peak groups.  $\dot{V}_{O_2}$ peak = peak oxygen uptake.



**Table 2.** Predictors of Death or Lung Transplantation Based on Random-Effects Cox Proportional Hazards Regression (Adjusted Analyses, 10-yr Follow-up)

	CPET-derived Variable in the Model						
	$\dot{V}O_2$ peak (% pred)	Wpeak (% pred)	$\dot{V}E/\dot{V}O_2$ peak	$\dot{V}E/\dot{V}CO_2$ peak	$\dot{V}E$ peak/MVVpred	SpO <sub>2</sub> peak (%)	$\Delta\dot{V}E/\Delta\dot{V}CO_2$ Slope
CPET-derived measure	0.964 (0.944–0.986); P = 0.001	0.969 (0.951–0.988); P = 0.001	1.085 (1.041–1.132); P < 0.001	1.060 (1.007–1.115); P = 0.025	0.780 (0.245–2.547); P = 0.693	0.973 (0.907–1.043); P = 0.442	1.020 (0.960–1.085); P = 0.521
Additional covariates in the model							
Sex	1.813 (1.062–3.010); P = 0.029	1.949 (1.1327–3.356); P = 0.016	1.651 (0.947–2.881); P = 0.077	1.829 (1.060–3.158); P = 0.030	1.787 (1.017–3.318); P = 0.043	1.619 (0.894–2.933); P = 0.112	1.724 (0.709–4.194); P = 0.230
CFRD	1.877 (0.831–4.242); P = 0.130	1.827 (0.802–4.159); P = 0.151	2.089 (0.913–4.782); P = 0.081	1.864 (0.831–3.704); P = 0.145	1.739 (0.778–3.888); P = 0.177	1.945 (0.830–4.557); P = 0.126	2.012 (0.607–6.667); P = 0.253
PA	2.624 (0.996–6.914); P = 0.051	2.928 (1.082–7.928); P = 0.034	2.429 (0.884–6.675); P = 0.085	2.925 (1.163–7.359); P = 0.023	2.842 (1.040–7.768); P = 0.042	2.778 (0.984–7.844); P = 0.054	1.766 (0.521–5.988); P = 0.361
Age at CPET, yr	0.933 (0.886–0.982); P = 0.008	0.940 (0.892–0.990); P = 0.019	0.932 (0.879–0.989); P = 0.019	0.941 (0.892–0.992); P = 0.024	0.924 (0.872–0.979); P = 0.007	0.938 (0.882–0.998); P = 0.045	0.945 (0.858–1.042); P = 0.256
FEV <sub>1</sub> % predicted	0.952 (0.932–0.973); P < 0.001	0.946 (0.927–0.966); P < 0.001	0.928 (0.910–0.946); P < 0.001	0.934 (0.918–0.951); P < 0.001	0.929 (0.907–0.950); P < 0.001	0.936 (0.915–0.957); P < 0.001	0.929 (0.899–0.960); P < 0.001
BMI (z-score)	1.078 (0.841–1.382); P = 0.553	0.991 (0.770–1.275); P = 0.942	1.022 (0.781–1.338); P = 0.873	1.046 (0.825–1.325); P = 0.713	0.958 (0.736–1.247); P = 0.750	0.968 (0.726–1.292); P = 0.828	1.038 (0.714–1.510); P = 0.845

*Definition of abbreviations:* BMI = body mass index; CFRD = cystic fibrosis-related diabetes; CPET = cardiopulmonary exercise testing; MVV = maximal voluntary ventilation; PA = chronic *Pseudomonas aeruginosa* infection; SpO<sub>2</sub>peak = oxygen saturation at peak exercise;  $\dot{V}E$ peak/MVVpred = breathing reserve index (MVV was calculated as FEV<sub>1</sub> × 35);  $\dot{V}E/\dot{V}CO_2$ peak = ventilatory equivalent for carbon dioxide;  $\dot{V}E/\dot{V}O_2$ peak = ventilatory equivalent for oxygen;  $\dot{V}O_2$ peak = peak oxygen consumption; Wpeak = peak work rate;  $\Delta\dot{V}E/\Delta\dot{V}CO_2$  slope = minute ventilation–carbon dioxide production relationship slope.  
Data are hazard ratios (95% confidence interval) and probability of a type I error.



**Figure 3.** Kaplan-Meier survival curve for four different patient groups identified using Ward hierarchical clustering.

parameters in a heterogeneous sample of both pediatric and adult patients with CF from Australia, Europe, and North America and a follow-up period of up to 14 years. Previous studies investigating the prognostic value of CPET-derived variables in patients with CF were all single-center studies, performed in either children (6, 7, 9) and both children and adults (4, 5, 8) with much fewer cases in some studies (9–61 deaths/lung transplants) compared with our study, and also shorter follow-up periods ranging from less than 1 year to 8 years (4–9). We extend the current knowledge on the predictive value of CPET-derived parameters in addition to established prognostic markers in a large and international cohort and identified  $\dot{V}_E/\dot{V}_{O_2}$  and  $\dot{V}_E/\dot{V}_{CO_2}$  as important prognostic factors.

In our study, patients with the highest  $\dot{V}_{O_2peak}$  ( $\geq 82\%$  predicted) had a 72% (HR, 0.278; 95% CI, 0.088–0.882) and 49% (HR, 0.507; 95% CI, 0.259–0.993) lower risk of dying or receiving a lung transplant in the following 10 years compared with patients in the middle (59–81% predicted) and lowest  $\dot{V}_{O_2peak}$  category ( $\leq 58\%$  predicted), respectively. These findings are in line with a landmark single-center study on 109 patients with CF (age 7–35 yr;  $FEV_1$  59% predicted [range, 24–95%]) demonstrating for the first time that a high aerobic exercise capacity is associated with lower risk of dying over 8 years in CF (4). We confirm the concept and extend these data in a large and international sample of pediatric and adult patients with CF treated in the

modern era. Indeed, in the study by Nixon and colleagues (4) more than 20 years ago, 8-year survival rates were much lower compared with our study population (56% vs. 85% survival rate). Although survival rates in CF have substantially improved over the last decades (16, 17), we also observed an overall higher  $\dot{V}_{O_2peak}$  in our study population compared with the Nixon study population (82% predicted [range, 23–151] vs. 70% predicted [range, 21–132]) using the same  $\dot{V}_{O_2peak}$  prediction equations (4). Our data clearly confirm that a well-preserved fitness remains important for survival in CF and possibly reflects improvements in treatment and the change in practice with regard to the acceptance of exercise as a key element in CF treatment and a stronger focus on exercise counseling.

In multivariable models, we identified  $\dot{V}_{O_2peak}$ ,  $W_{peak}$ , and  $\dot{V}_E/\dot{V}_{O_2}$  and  $\dot{V}_E/\dot{V}_{CO_2}$  at peak exercise as predictors of death/LTx even after rigorous adjustment for established predictors of death in CF (1–3). In our study, age at CPET and  $FEV_1$  were significant predictors of death/LTx in all fully adjusted models. Lung disease severity, assessed with  $FEV_1$ , is well known to be a strong predictor of survival in CF (18) and whether  $FEV_1$  or  $\dot{V}_{O_2peak}$  is a stronger predictor of survival in CF is controversial (4–6). Most (5–8), but not all, studies (4) evaluating the prognostic value of CPET-derived parameters showed strong associations of  $FEV_1$  with mortality in multivariable models. In our study,  $FEV_1$  explained the greatest variance of

subsequent death/LTx in the adjusted Cox models and  $FEV_1$  was the variable with the highest importance for phenotypic clustering (Figure 3). Interestingly, all CPET-related variables (except  $W_{peak}$ ) had a higher importance for the outcome death/LTx than other established predictors of survival in CF. This finding supports the conclusion that CPET-derived variables are important prognostic factors in CF.

Beside  $\dot{V}_{O_2peak}$  and  $W_{peak}$  as known predictors of mortality (4–6, 9),  $\dot{V}_E/\dot{V}_{O_2}$  and  $\dot{V}_E/\dot{V}_{CO_2}$  at peak exercise were also associated with death/LTx in our patient population. The  $\dot{V}_E/\dot{V}_{O_2peak}$  (a marker of ventilatory efficiency at peak exercise) was previously identified as stronger predictor of survival than  $\dot{V}_{O_2peak}$  in univariate analyses of adult patient data (5). In another study using multivariable analysis in a cohort of 127 children with CF aged 11–14 years (7), the equation to predict a higher risk of mortality included  $\dot{V}_E/\dot{V}_{O_2peak}$ , BMI, and  $FEV_1$  (%). These data are supported by our cluster analysis (see variable importance plot in Figure 3), indicating that  $\dot{V}_E/\dot{V}_{O_2peak}$  has a greater importance for phenotype clustering than  $\dot{V}_{O_2peak}$ . The highest values for  $\dot{V}_E/\dot{V}_{O_2peak}$  and  $\dot{V}_E/\dot{V}_{CO_2peak}$  were observed in the high-risk group (cluster 3 in Table 3), which was also the group with the lowest FVC that is suggestive of a higher dead space ventilation and reduced breathing reserve (i.e.,  $\dot{V}_{Epeak}/MVV_{pred}$  was highest in cluster 3). Interestingly, clusters 3 and 4 were both characterized by severe lung function impairment, but patients in cluster 3 were younger, had worse nutritional status (BMI), and substantially lower exercise capacity, whereas CFTR genotype showed no obvious differences compared with the other clusters (e.g., proportion of patients with severe CFTR mutations). Thus, we can only speculate on underlying reasons, such as lower daily physical activity levels, but these data were not available. Nevertheless, this cluster of patients may especially benefit from regular monitoring of exercise capacity and appropriate exercise counseling to improve nutritional status, one of the key predictors of aerobic exercise capacity in CF (19).

This study has limitations. We included data from 10 CF centers from different countries with likely heterogeneous diagnostic and treatment regimens possibly

**Table 3.** Cystic Fibrosis Risk Groups Defined by Clinical and Physiologic Parameters Using Ward Hierarchical Clustering

	Cluster 1 (n = 207)	Cluster 2 (n = 130)	Cluster 3 (n = 54)	Cluster 4 (n = 33)
Exposure variable				
Female sex	85 (41)	55 (42)	28 (52)	11 (33)
Age at CPET, yr	15.9 ± 5.1*	14.5 ± 3.5 <sup>††</sup>	17.1 ± 4.2 <sup>†§</sup>	29.9 ± 6.1 <sup>**§</sup>
BMI (z-score)	-0.17 ± 0.8 <sup>*  ¶</sup>	-1.03 ± 0.8 <sup>†  </sup>	-1.86 ± 1.1 <sup>†§¶</sup>	-0.75 ± 0.8 <sup>*§</sup>
Genotype, n (%)**				
CFTR, both alleles from classes I-III	154 (74.4)	97 (74.6)	38 (70.4)	19 (57.6)
CFTR, at least one allele from classes IV-V	14 (6.8)	6 (4.6)	1 (1.9)	0 (0)
CFTR, at least one allele unknown/not available	39 (18.8)	27 (20.8)	15 (27.7)	14 (42.4)
CFRD	11 (5)*	10 (8) <sup>‡</sup>	3 (6) <sup>§</sup>	9 (27) <sup>*‡§</sup>
Chronic PA colonization	120 (58) <sup>*  ¶</sup>	97 (75) <sup>  </sup>	42 (78) <sup>¶</sup>	30 (91)*
Pancreatic insufficiency	188 (91) <sup>*  ¶</sup>	122 (94) <sup>  </sup>	53 (98) <sup>¶</sup>	32 (97)*
Lung function				
FEV <sub>1</sub> % predicted	87.6 ± 13.2 <sup>*  ¶</sup>	69.2 ± 16.4 <sup>††  </sup>	43.1 ± 13.9 <sup>†¶</sup>	49.3 ± 13.9 <sup>*‡</sup>
FVC % predicted	95.2 ± 11.1 <sup>*  ¶</sup>	80.9 ± 12.6 <sup>††  </sup>	60.9 ± 15.2 <sup>†§¶</sup>	71.9 ± 15.6 <sup>*‡§</sup>
CPET variables				
V <sub>O<sub>2</sub>peak</sub> % predicted	79.4 ± 16.2 <sup>*  ¶</sup>	64.8 ± 15.2 <sup>†  </sup>	58.8 ± 15.3 <sup>†§¶</sup>	73.4 ± 16.9 <sup>*§</sup>
W <sub>peak</sub> % predicted	103.9 ± 20.1 <sup>*  ¶</sup>	83.2 ± 15.4 <sup>†  </sup>	66.4 ± 21.3 <sup>†§¶</sup>	84.3 ± 21.1 <sup>*§</sup>
V <sub>E</sub> /V <sub>O<sub>2</sub>peak</sub>	37.16 ± 5.6 <sup>  ¶</sup>	40.22 ± 6.5 <sup>†  </sup>	48.01 ± 8.2 <sup>†§¶</sup>	37.70 ± 6.9 <sup>¶</sup>
V <sub>E</sub> /V <sub>CO<sub>2</sub>peak</sub>	31.75 ± 4.5 <sup>  ¶</sup>	34.86 ± 5.3 <sup>†  </sup>	41.58 ± 7.7 <sup>†§¶</sup>	32.50 ± 4.2 <sup>¶</sup>
Sp <sub>O<sub>2</sub>peak</sub> , %	96.4 ± 3.1 <sup>*¶</sup>	95.3 ± 3.9 <sup>†‡</sup>	91.5 ± 5.4 <sup>†¶</sup>	91.1 ± 4.8 <sup>*‡</sup>
V <sub>E</sub> peak/MVVpred	0.8 ± 0.2 <sup>*†  </sup>	0.9 ± 0.2 <sup>†  </sup>	1.2 ± 0.2 <sup>†§¶</sup>	1.1 ± 0.3 <sup>*§</sup>
ΔV <sub>E</sub> /ΔV <sub>CO<sub>2</sub></sub> slope	28.9 ± 4.1 <sup>  ¶</sup>	31.4 ± 4.3 <sup>*</sup>	34.4 ± 8.6 <sup>†¶</sup>	29.4 ± 3.7
Outcomes death or LTx				
Up to 5-yr follow-up	1 (0) <sup>*¶</sup>	4 (3) <sup>†</sup>	19 (35) <sup>†¶</sup>	4 (12)*
Up to 10-yr follow-up	3 (1) <sup>*  ¶</sup>	17 (13) <sup>†  </sup>	34 (63) <sup>†§¶</sup>	9 (27) <sup>*§</sup>
Entire study period	4 (2) <sup>*  ¶</sup>	19 (15) <sup>†‡  </sup>	36 (67) <sup>†¶</sup>	13 (39) <sup>*‡</sup>

*Definition of abbreviations:* BMI = body mass index; CFRD = cystic fibrosis–related diabetes; CFTR = cystic fibrosis transmembrane conductance regulator; CPET = cardiopulmonary exercise testing; LTx = lung transplantation; MVV = maximal voluntary ventilation; PA = chronic *Pseudomonas aeruginosa* infection; Sp<sub>O<sub>2</sub>peak</sub> = oxygen saturation at peak exercise; V<sub>E</sub>peak/MVVpred = breathing reserve index (MVV was calculated as FEV<sub>1</sub> × 35); V<sub>E</sub>/V<sub>CO<sub>2</sub>peak</sub> = ventilatory equivalent for carbon dioxide; V<sub>E</sub>/V<sub>O<sub>2</sub>peak</sub> = ventilatory equivalent for oxygen; V<sub>O<sub>2</sub>peak</sub> = peak oxygen consumption; W<sub>peak</sub> = peak work rate.

Data are mean ± SD or number (%). Data for ΔV<sub>E</sub>/ΔV<sub>CO<sub>2</sub></sub> slope were only available for a subset of patients (cluster 1 = 119, cluster 2 = 56, cluster 3 = 18, and cluster 4 = 21).

Differences among clusters were analyzed using chi-square tests for categorical variables and Kruskal-Wallis, as appropriate. The Bonferroni-corrected significance level for these tests was 0.008 (overall significance level [0.05] divided by number of tests, which was six because we compared four clusters). If the test passed the significance level, this is indicated by a footnote as shown below.

\*Difference between cluster 1 and 4.

†Difference between cluster 2 and 3.

‡Difference between cluster 2 and 4.

§Difference between cluster 3 and 4.

||Difference between cluster 1 and 2.

¶Difference between cluster 1 and 3.

\*\*Differences between clusters based on chi-square statistics ( $P = 0.039$ ). If chi-square statistics are calculated for patients with known genotype by excluding patients with at least one “unknown; nonavailable” CFTR allele, no difference in genotype is observed between the groups ( $P = 0.338$ ).

introducing bias. For these reasons, all statistical models were adjusted for study center as random intercept. The models incorporate the fact that measurements from the same center are not independent, but more similar to each other by assigning an individual intercept for each center. This way, CF care centers with systematically biased values (e.g., clearly above or below the average) should not substantially change the overall results.

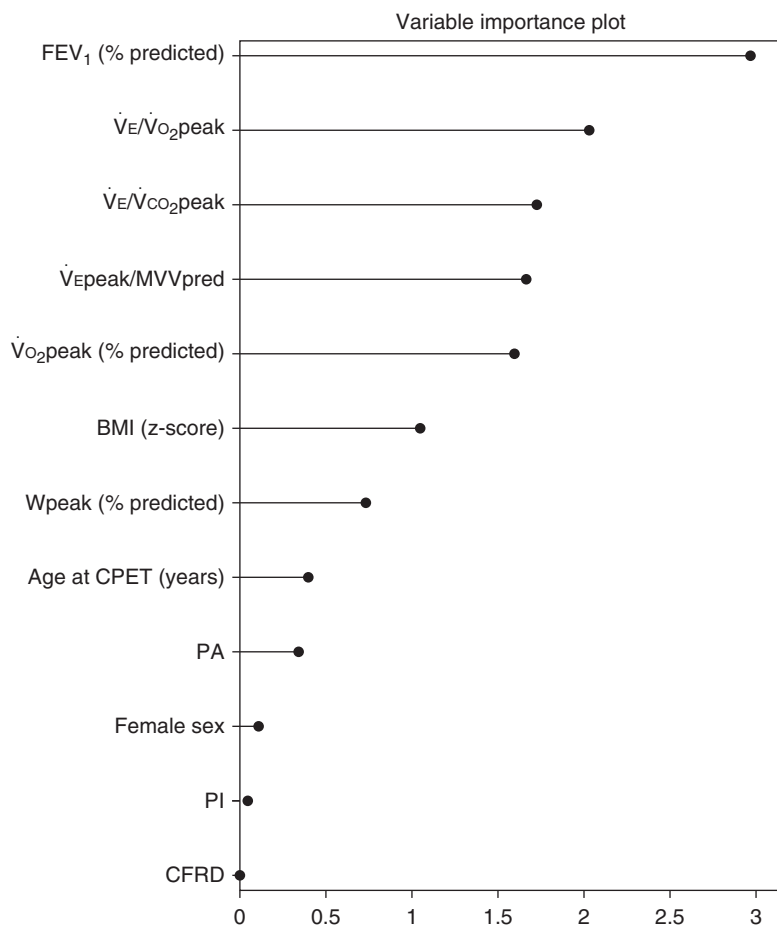
*A priori*, we included data from patients for whom follow-up information on survival and LTx was available greater than or equal to 5 years after CPET. Thus,

patients who left their respective center shortly after the CPET were excluded already when the data were collected. This approach might have introduced some bias. It is possible that we excluded more healthy and active patients who left their hometown for education or work. However, it is also possible that we excluded patients who were not doing well and were unhappy with their center care. By excluding the few cases on whom follow-up data were not available for at least 5 years ( $n = 40$ ), the number of cases available for analysis was somewhat reduced. However, because clear information on survival or death/LTx was

available on all cases for at least 5 years, the analyses are more sound.

Because of a large number of missing data on both CFTR mutations (22%) and unequal distribution among cases and survivors, we were not able to adjust the models for CFTR genotype, a well-known predictor of survival in CF (20, 21).

The data used for this study were gathered retrospectively. Thus, we had to limit the variables used as predictors of prognosis to those that were commonly assessed during the years 2000–2007. Because, for example, colonization with mycobacteria other than tuberculosis or



**Figure 4.** Variable importance plot generated by the unsupervised random forest algorithm from the 12 clinical and physiologic parameters used for clustering. Higher values corresponded to higher importance of the variable for clustering. BMI = body mass index; CFRD = cystic fibrosis–related diabetes; CPET = cardiopulmonary exercise testing; MVV = maximal voluntary ventilation; PA = chronic *Pseudomonas aeruginosa* infection; PI = pancreatic insufficiency;  $\dot{V}_{Epeak}/MVV_{pred}$  = breathing reserve index (MVV was calculated as FEV<sub>1</sub> × 35);  $\dot{V}_E/\dot{V}_{CO_2}$ peak = ventilatory equivalent for carbon dioxide;  $\dot{V}_E/\dot{V}_{O_2}$ peak = ventilatory equivalent for oxygen;  $\dot{V}_{O_2}$ peak = peak oxygen consumption; Wpeak = peak work rate.

*B. cepacia* complex was not vigorously assessed at that time, such information could not be included in the models.

It is important to note that transplant predictions (i.e., timing for referral for LTx) are usually based on 2-year mortality data in CF (22). Our study was not designed to assess short-term outcomes and established predictors of mortality for short-term prediction of mortality, such as 6-minute-walk test distance, and infection with mycobacteria or *B. cepacia* complex (22) were not validly available in our retrospective data collection. In an exploratory analysis restricted to 2-year follow-up data ( $n = 433$ ), which includes only a small number of death/LTx cases

( $n = 11$ ), all CPET parameters except the  $\Delta\dot{V}_E/\Delta\dot{V}_{CO_2}$  slope were significantly associated with death/LTx. However, these findings must be interpreted with caution and the short-term prognostic value of CPET-derived variables should be investigated in a well-designed prospective study including known predictors of referral for LTx (22).

Finally, our data were collected before initiation of CFTR modulator therapies. Because these drugs may impact on lung function and overall survival, the predictors should be reevaluated in the context of highly effective CFTR modulators that may become available to most patients with CF in the upcoming years.

Because aerobic fitness and ventilatory efficiency measured during CPET are related to prognosis, it is important to show whether improvements in these measures would translate into increased survival. Several randomized controlled trials have shown that exercise capacity can be increased in CF by exercise conditioning programs (23). However, no conditioning studies with survival/mortality being the primary outcome are available to date. Because information from exercise testing in CF is relevant to patient care in several aspects, such as detecting adverse reactions to exercise, understanding exercise limitations, guiding a conditioning program, or motivating people to engage in strenuous physical activities, an increasing number of centers are including formal exercise testing in their routine work-up of patients with CF. Including the information from these tests in patient registries might in the future show whether patients who get fitter will benefit with respect to longevity.

## Conclusions

Data collected during CPET, such as  $\dot{V}_{O_2}$ peak, Wpeak,  $\dot{V}_E/\dot{V}_{O_2}$ peak, and  $\dot{V}_E/\dot{V}_{CO_2}$ peak, provide prognostic information in addition to established predictors of death/LTx in CF. Cluster analysis revealed that CPET-derived measures were important for phenotyping. The phenotype with the highest risk for death/LTx identified by Ward hierarchical clustering is characterized by poor lung function, nutritional status, and exercise capacity and may thus especially benefit from regular monitoring of exercise capacity, nutritional counseling, and exercise intervention. ■

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

**Contributors of the Prognostic Value of CPET in CF Study Group:** Sibylle Junge, Clinic for Pediatric Pneumology and Neonatology, Hannover Medical School, Hannover, Germany; Christine Smaczny, Medical Clinic I, Pneumology and Allergology, University Hospital Frankfurt/Main, Goethe University, Frankfurt/Main, Germany; Sarah Rand, Specialist Pediatric Cystic Fibrosis Physiotherapist, Great Ormond Street Hospital, London, and Senior Teaching Fellow, Institute of Child Health, University College London, United Kingdom; and Charlotte Dawson, Cystic Fibrosis Nurse Specialist, Great Ormond Street Hospital, London, United Kingdom.



## References

- Hayllar KM, Williams SG, Wise AE, Pouria S, Lombard M, Hodson ME, *et al.* A prognostic model for the prediction of survival in cystic fibrosis. *Thorax* 1997;52:313–317.
- Buzzetti R, Salvatore D, Baldo E, Forneris MP, Lucidi V, Manunza D, *et al.* An overview of international literature from cystic fibrosis registries: 1. Mortality and survival studies in cystic fibrosis. *J Cyst Fibros* 2009;8:229–237.
- Vieni G, Faraci S, Collura M, Lombardo M, Traverso G, Cristadoro S, *et al.* Stunting is an independent predictor of mortality in patients with cystic fibrosis. *Clin Nutr* 2013;32:382–385.
- Nixon PA, Orenstein DM, Kelsey SF, Doershuk CF. The prognostic value of exercise testing in patients with cystic fibrosis. *N Engl J Med* 1992; 327:1785–1788.
- Moorcroft AJ, Dodd ME, Webb AK. Exercise testing and prognosis in adult cystic fibrosis. *Thorax* 1997;52:291–293.
- Pianosi P, Leblanc J, Almudevar A. Peak oxygen uptake and mortality in children with cystic fibrosis. *Thorax* 2005;60:50–54.
- Hulzebos EH, Bomhof-Roordink H, van de Weert-van Leeuwen PB, Twisk JW, Arets HG, van der Ent CK, *et al.* Prediction of mortality in adolescents with cystic fibrosis. *Med Sci Sports Exerc* 2014;46: 2047–2052.
- Tantisira KG, Systrom DM, Ginns LC. An elevated breathing reserve index at the lactate threshold is a predictor of mortality in patients with cystic fibrosis awaiting lung transplantation. *Am J Respir Crit Care Med* 2002;165:1629–1633.
- van de Weert-van Leeuwen PB, Sliker MG, Hulzebos HJ, Kruitwagen CL, van der Ent CK, Arets HG. Chronic infection and inflammation affect exercise capacity in cystic fibrosis. *Eur Respir J* 2012;39: 893–898.
- Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 1963;58:236–244.
- Godfrey S, Davies CT, Wozniak E, Barnes CA. Cardio-respiratory response to exercise in normal children. *Clin Sci* 1971;40:419–431.
- Orenstein DM. Assessment of exercise pulmonary function. In: Rowland TW, editor. Pediatric laboratory exercise testing clinical guidelines. Champaign, IL: Human Kinetics Publishers; 1993. pp. 141–163.
- Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, *et al.*; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012;40:1324–1343.
- Rosario AS, Kurth BM, Stolzenberg H, Ellert U, Neuhauser H. Body mass index percentiles for children and adolescents in Germany based on a nationally representative sample (KiGGS 2003–2006). *Eur J Clin Nutr* 2010;64:341–349.
- Hemmelmann C, Brose S, Vens M, Hebebrand J, Ziegler A. [Percentiles of body mass index of 18–80-year-old German adults based on data from the Second National Nutrition Survey]. *Dtsch Med Wochenschr* 2010;135:848–852.
- MacKenzie T, Gifford AH, Sabadosa KA, Quinton HB, Knapp EA, Goss CH, *et al.* Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the Cystic Fibrosis Foundation patient registry. *Ann Intern Med* 2014;161:233–241.
- George PM, Banya W, Pareek N, Bilton D, Cullinan P, Hodson ME, *et al.* Improved survival at low lung function in cystic fibrosis: cohort study from 1990 to 2007. *BMJ* 2011;342:d1008.
- Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. *N Engl J Med* 1992;326: 1187–1191.
- Radtke T, Hebestreit H, Gallati S, Schneiderman JE, Braun J, Stevens D, *et al.*; CFTR-Exercise study group. CFTR genotype and maximal exercise capacity in cystic fibrosis: a cross-sectional study. *Ann Am Thorac Soc* 2018;15:209–216.
- McKone EF, Goss CH, Aitken ML. CFTR genotype as a predictor of prognosis in cystic fibrosis. *Chest* 2006;130:1441–1447.
- de Gracia J, Mata F, Alvarez A, Casals T, Gatner S, Vendrell M, *et al.* Genotype-phenotype correlation for pulmonary function in cystic fibrosis. *Thorax* 2005;60:558–563.
- Weill D, Benden C, Corris PA, Dark JH, Davis RD, Keshavjee S, *et al.* A consensus document for the selection of lung transplant candidates: 2014—an update from the Pulmonary Transplantation Council of the International Society for Heart and Lung Transplantation. *J Heart Lung Transplant* 2015;34:1–15.
- Radtke T, Nevitt SJ, Hebestreit H, Kriemler S. Physical exercise training for cystic fibrosis. *Cochrane Database Syst Rev* 2017;11: CD002768.