Cardiopulmonary exercise testing provides additional prognostic information in cystic fibrosis

Hebestreit, Helge ; Hulzebos, Erik HJ ; Schneiderman, Jane E ; Karila, Chantal ; Boas, Steven R ; Kriemler, Susi ; Dwyer, Tiffany ; Sahlberg, Margareta ; Urquhart, Don S ; Lands, Larry C ; Ratjen, Felix ; Takken, Tim ; Varanistkaya, Liobou ; Rücker, Viktoria ; Hebestreit, Alexandra ; Usemann, Jakob ; Radtke, Thomas

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 (VO2peak) following rigorous adjustment for other predictors. MEASUREMENTS AND MAIN RESULTS: Data from 10 CF-centers in Australia, Europe and North America were collected retrospectively. 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 (LTx) was analyzed using Cox proportional hazards regression. In addition, phenotyping using hirarchical Ward's clustering was performed to characterize high risk subgroups. Cox regression showed - even after adjustment for sex, forced expiratory volume in 1s (%predicted), body mass index (z-score), age at CPET, Pseudomonas aeruginosa status, and CF-related diabetes as covariates in the model - that VO2peak in %predicted, hazard ratio (HR) 0.964 [95%-CI: 0.944-0.986], peak work rate (%predicted, HR 0.969 [0.951-0.988], ventilatory equivalent for oxygen (VE/VO2peak) HR 1.085 [1.041-1.132], and carbon dioxide (VE/VCO2peak), HR 1.060 [1.007-1.115], all $P<0.05$) were significant predictors of death or LTx at 10 years 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: In conclusion, CPET provides additional prognostic information to established predictors of death/LTx in CF. High risk patients may especially benefit from regular monitoring of exercise capacity and exercise counselling.

DOI: https://doi.org/10.1164/rccm.201806-1110oc

DOI: https://doi.org/10.1164/rccm.201806-1110oc

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Title: Cardiopulmonary exercise testing provides additional prognostic information in cystic fibrosis

Authors: Helge Hebestreit[1], Erik H.J. Hulzebos[2], Jane E. Schneiderman[3], Chantal Karila[4], Steven R. Boas[5], Susi Kriemler[6], Tiffany Dwyer[7], Margareta Sahlberg[8], Don S. Urquhart[9], Larry C Lands[10], Felix Ratjen[3], Tim Takken[2], Liobou Varanistkaya[1], Viktoria Rücker[11], Alexandra Hebestreit[1], Jakob Usemann[12], Thomas Radtke[6] for the Prognostic value of CPET in CF study group*

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

Corresponding author: Prof. Dr. Helge Hebestreit; Universitäts-Kinderklinik; Josef-Schneider-Str. 2; 97080 Würzburg, Germany; E-mail: hebestreit@uni-wuerzburg.de

Author contributions: Conception and design (HH); Acquisition of data (AH, DSU, CK, EHJH, HH, JES, LCL, LV, MS, SK, SRB, TT, TD); Genotype classification (TR, HH); Statistical analysis (JU, TR, VR); Interpretation (FR, HH, JU, TR); First draft (HH, TR); All authors edited, reviewed, and approved the final version of the manuscript.

Running title: Cardiopulmonary exercise testing and prognosis in cystic fibrosis

Descriptor number: 9.17: Cystic Fibrosis: Translational & Clinical Studies; 8.22 Peripheral Muscle Function In Lung Disease; 8.13 Exercise in Health & Disease

Total word count: 3527

At a Glance Commentary

Scientific Knowledge on the Subject: Cardiopulmonary exercise testing (CPET) 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 CPET 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 CPET parameters and identifies the ventilatory equivalent for oxygen (VE/VO_2) and carbon dioxide (VE/VCO_2) at peak exercise as important predictors of the compound outcome survival/lung transplantation (LTx). 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.

***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 Paediatric Cystic Fibrosis Physiotherapist, Great Ormond Street Hospital, London, WC1N 3JH; and Senior Teaching Fellow, Institute of Child Health, University College London, WC1N 1EH.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Charlotte Dawson, Cystic Fibrosis Nurse Specialist, Great Ormond Street Hospital, London, WC1N 3JH.

"This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org

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 (VO_2peak) following rigorous adjustment for other predictors.

Measurements and Main Results: Data from 10 CF-centers in Australia, Europe and North America were collected retrospectively. 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 (LTx) was analyzed using Cox proportional hazards regression. In addition, phenotyping using hirarchical Ward's clustering was performed to characterize high risk subgroups. Cox regression showed - even after adjustment for sex, forced expiratory volume in 1s (%predicted), body mass index (z-score), age at CPET, *Pseudomonas aeruginosa* status, and CF-related diabetes as covariates in the model - that VO_2peak in %predicted, hazard ratio (HR) 0.964 [95%-CI: 0.944-0.986], peak work rate (%predicted, HR 0.969 [0.951-0.988], ventilatory equivalent for oxygen (VE/VO_2peak) HR 1.085 [1.041-1.132], and carbon dioxide (VE/VCO_2peak), HR 1.060 [1.007-1.115], all $P<0.05$) were significant predictors of death or LTx at 10 years 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:** In conclusion, CPET provides additional prognostic information to established predictors of death/LTx in CF. High risk patients may especially benefit from regular monitoring of exercise capacity and exercise counselling.

Abstract Word Count: 250

Keywords: cystic fibrosis, prognosis, survival, peak oxygen uptake, exercise testing, lung transplantation

Introduction

In patients with cystic fibrosis (CF), an increased risk of death has been associated with poor values of forced expiratory volume in 1s (FEV_1), short stature, low body mass index (BMI), pancreatic insufficiency, CF-related diabetes (CFRD), and chronic infection with specific pathogens, e.g., *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). Peak oxygen uptake (VO_2peak) and change in VO_2peak over time, peak work rate (Wpeak), the respiratory equivalent for oxygen (minute ventilation (VE) divided by VO_2) at peak exercise (VE/VO_2peak) and the breathing reserve index at the ventilatory threshold (VE 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 due to 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 LTx (e.g., FEV_1 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 as numerous other factors are associated with the outcome.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Another possibility of identifying high risk patients is to use cluster analysis, for example Ward's hierachical clustering – an unbiased, data-driven approach to define clinical/physiological 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 VO_2peak (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 years after CPET to allow comparisons to 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 ≥10 years who had a full CPET meeting prespecified criteria between 1st January 2000 and 31st December 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 supplementary material).

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 (Table E1) and criteria for a maximal effort are available in the online supplement.

Statistical analysis

Data are presented as N (%), mean±SD (ranges) and hazard ratios with 95% confidence intervals (CI). Data from lung function testing and CPET was 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 VO_2peak (and Wpeak) with LTx and/or death was visualized by plotting survival (Kaplan-Meier curves) of three groups using cut-offs employed by Nixon et al. (4) for Cox proportional hazards regression. The Cox proportional hazards assumption was verified graphically (log-log plot) and tested using Schoenfeld's 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 an 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 (VO_2peak, Wpeak, VEpeak/MVVpred, VE/VO_2peak, VE/VCO_2peak, SpO_2peak, and the $\Delta VE/\Delta VCO_2$ slope) in addition to known predictors of survival in CF in multivariable models. Cox proportional hazards regression models were performed including FEV_1 (% predicted), BMI (z-score), age at CPET, sex, and the binary coded variables (yes/no) chronic *Pseudomonas aeruginosa* colonization, and CF-related diabetes (CFRD). Exocrine pancreatic insufficiency was a pre-defined covariate to be considered for the statistical model. However, since none of the cases (LTx/death) in

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

our dataset was pancreatic sufficient, we were not able to compute effect estimates and 95% CI's. Consequently, pancreatic insufficiency was not included in the statistical models. Due to the fact that data on Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) genotype was 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 controls 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 due to low numbers of cases (N=8) from only three study centers with >10-year follow-up data. In addition, we evaluated predictors of survival for different time periods (e.g., 5, 8 and 10 years) using Cox models with time-varying covariates.

In addition, an unsupervised, data-driven approach was employed to explore the relevance of CPET-derived parameters to predict death/LTx in CF patients. Physiological and clinical parameters were used as input to define clusters by Ward's 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. 2012, College Station, Texas, USA and R version 3.0.2 (http://www.R-project.org). A probability for a type I error $P<0.05$ 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. 433 of 510 CPET's fulfilled at least one of the four predefined criteria

for a maximal effort (i.e., plateauing of VO_2, or RER, heart rate or VE exceeding the above thresholds) and were, thus, used for the final analysis. Patients with a non-maximal effort during CPET (N=77) were younger (13.7±3.5 years, $P<0.001$), had comparable FEV_1 (74.0±21.7 % predicted, $P=0.868$) and showed a trend of lower VO_2peak (77.3±18.6 %predicted, $P=0.066$) compared to the final study population. Of 77 excluded patients, 8 (10%) died or had a LTx during the entire follow-up period. We noticed differences in deaths/LTx among study centers (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 Table E3 and E4 in the online supplements.

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 VO_2peak categories are shown in the Kaplan-Meier curve (Figure 2). A figure with Wpeak categories is shown in the online supplement (Figure E1).

Results of Cox proportional hazard models (univariate and adjusted analyses)

In univariable random effects models with study center adjustment, the CPET-derived information on VO_2peak, Wpeak, VE/VO_2peak, VE/VCO_2peak, VEpeak/MVVpred, SpO_2peak and the $\Delta VE/\Delta VCO_2$ slope was each a significant predictor of death or LTx (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 Tables E6). A low FEV_1 and BMI, as well as chronic *Pseudomonas aeruginosa* colonization were also related to death or LTx (Tables E5-E6).

A subanalysis restricted to patients with advanced lung disease ($FEV_1 \leq 40\%$ predicted) identified age at CPET, VO_2peak, Wpeak, VE/VO_2peak, and VE/VCO_2peak as significant predictors of death or LTx (Table E7). Comparable results for CPET-variables were found in non center-adjusted analyses (Table E8).

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

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), VO_2peak, Wpeak, VE/VO_2peak, VE/VCO_2peak, VEpeak/MVVpred, and SpO_2peak as well as FEV_1 and BMI were each significantly associated with death or LTx (Tables E9 and E10). Results from these short-term models were similar to complete dataset models (Tables E5-6), except for a significant effect of chronic *Pseudomonas aeruginosa* colonization on death or LTx (Tables E9-10).

When adjusting the models for sex, age at CPET, FEV_1 (% predicted), BMI (z-score), chronic *Pseudomonas aeruginosa* colonization, and CFRD the variables VO_2peak, Wpeak, VE/VO_2peak and VE/VCO_2peak remained significant predictors of death or LTx (Table 2). Characteristics of patients within the three VO_2peak categories are shown in Table E11. Patients in the highest and middle VO_2peak groups (≥82% predicted and 59-81% predicted, respectively) had a 72% (HR 0.278, 95% CI 0.088 to 0.882, *P*=0.030) and 49% (HR 0.507, 95% CI 0.259 to 0.993, *P*=0.048) lower risk of dying or to receive a lung transplant within the next 10 years compared to patients with a VO_2peak of ≤58% of the predicted value. In the model with Wpeak, patients in the highest Wpeak group (≥92% predicted) showed a trend for lower chance for death or LTx (HR 0.417, 95% CI 0.155 to 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 years, data not shown).

Results of Ward's hierarchical clustering

All CPET parameters and FEV_1 had a high variable importance to define clusters, as indicated by the Forest plot (Figure 3). However, the binary coded variables (sex, chronic *Pseudomonas aeruginosa* infection, pancreatic insufficiency, and CFRD) were less important. Based on the forest plot, the eight continuous variables (FEV_1, BMI, Wpeak, age at CPET, VO_2peak, VEpeak/MVVpred, VE/VO_2peak,

and VE/VCO_2peak) were introduced into the PCA, and five orthogonal factors explained more than 95% of variance (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 since the prevalence of the primary outcome death/LTx during the ten-year study period (63%) was highest compared to 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 *Pseudomonas* aeruginosa colonization, low BMI z-score, and low FEV_1 and FVC values (Figure E2).

We further identified a group of older patients (N=33, cluster 4) with poor lung function, high prevalence of chronic *Pseudomonas* 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 to 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 VO_2peak and Wpeak as key predictors of survival, but also identifies other CPET measures (VE/VO_2peak and VE/VCO_2peak) that may be of great prognostic significance. The use of cluster analysis further suggests ´at risk' phenotypes in whom early recognition, nutritional counselling and exercise intervention could be most beneficial.

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

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

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 to 61 deaths/lung transplants) compared to 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 VE/VO_2 and VE/VCO_2 as important prognostic factors.

In our study, patients with the highest VO_2peak (≥82% predicted) had a 72% (HR 0.278, 95% CI 0.088 to 0.882) and 49% (HR 0.507, 95% CI 0.259 to 0.993) lower risk of dying or receiving a lung transplant in the following 10 years compared to patients in the middle (59 to 81% predicted) and lowest VO_2peak category (≤58% predicted), respectively. These findings are in line with a landmark single-center study on 109 patients with CF [(age 7-35 years, 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 et al. (4) more than 20 years ago, 8-year survival rates were much lower compared with our study population (56% versus 85% survival rate). While survival rates in CF have substantially improved over the last decades (16, 17), we also observed an overall higher VO_2peak in our study population compared to the Nixon study population [(82 % predicted (range 23-151) versus 70 % predicted (range 21 to 132)] using the same VO_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 as well as the change in practice with regard to the acceptance of exercise as a key element in CF treatment and a stronger focus on exercise counselling.

In multivariable models, we identified VO_2peak, Wpeak, and VE/VO_2 and VE/VCO_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 VO_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 Wpeak) 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 VO_2peak and Wpeak as know predictors of mortality (4-6, 9), VE/VO_2 and VE/VCO_2 at peak exercise were also associated with death/LTx in our patient population. The VE/VO_2peak – a marker of ventilatory efficiency at peak exercise - was previously identified as stronger predictor of survival than VO_2peak in univariate analyses of adult patient data (5). In another study employing 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 VE/VO_2peak, BMI and FEV_1 (%). These data are supported by our cluster analysis (see variable importance plot, Figure 3), indicating that VE/VO_2peak has a greater importance for phenotype clustering than VO_2peak. The highest values for VE/VO_2peak and VE/VO_2peak were observed in the high-risk group (cluster 3, 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., VEpeak/MVVpred 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, while CFTR genotype showed no obvious differences compared to the other clusters (e.g., proportion of patients with severe CFTR mutations). Thus, we can

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

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 counselling 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 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 ≥ 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. On the other hand it is also possible that we excluded patients who were not doing well and 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, since clear information on survival or death/LTx was available on all cases for at least 5 years, the analyses are more sound.

Due to 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.

Since for example colonization with Mycobacteria other than tuberculosis or Burkolderia 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 Burkholderia 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 $\Delta VE/\Delta VCO_2$ 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. Since these drugs may impact on lung function and overall survival, the predictors should be re-evaluated in the context of highly effective CFTR modulators that may become available to the majority of CF patients in the upcoming years.

As 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. Since 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

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

activities, an increasing number of centers are including formal exercise testing in their routine work-up of patients with CF. By including the information from these tests into patient registries might - in the future - allow to determine whether patients who get fitter will benefit with respect to longevity.

Conclusions

Data collected during CPET such as VO_2peak, Wpeak, VE/VO_2peak, and VE/VCO_2peak 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 Wards hirachical clustering is characterized by poor lung function, nutritional status and exercise capacity and may thus especially benefit from regular monitoring of exercise capacity, nutritional counselling, and exercise intervention.

References

1. Hayllar KM, Williams SG, Wise AE, Pouria S, Lombard M, Hodson ME, Westaby D. A prognostic model for the prediction of survival in cystic fibrosis. *Thorax* 1997; 52: 313-317.

2. Buzzetti R, Salvatore D, Baldo E, Forneris MP, Lucidi V, Manunza D, Marinelli I, Messore B, Neri AS, Raia V, Furnari ML, Mastella G. An overview of international literature from cystic

fibrosis registries: 1. Mortality and survival studies in cystic fibrosis. *J Cyst Fibros* 2009; 8: 229-237.

3. Vieni G, Faraci S, Collura M, Lombardo M, Traverso G, Cristadoro S, Termini L, Lucanto MC, Furnari ML, Trimarchi G, Triglia MR, Costa S, Pellegrino S, Magazzu G. Stunting is an independent predictor of mortality in patients with cystic fibrosis. *Clinical Nutrition* 2013; 32: 382-385.

4. Nixon PA, Orenstein DM, Kelsey SF, Doershuk CF. The prognostic value of exercise testing in patients with cystic fibrosis. *N Eng J Med* 1992; 327: 1785-1788.

5. Moorcroft AJ, Dodd ME, Webb AK. Exercise testing and prognosis in adult cystic fibrosis. *Thorax* 1997; 52: 291-293.

6. Pianosi P, Leblanc J, Almudevar A. Peak oxygen uptake and mortality in children with cystic fibrosis. *Thorax* 2005; 60: 50-54.

7. Hulzebos EH, Bomhof-Roordink H, van de Weert-van Leeuwen PB, Twisk JW, Arets HG, van der Ent CK, Takken T. Prediction of mortality in adolescents with cystic fibrosis. *Med Sci Sports Exerc* 2014; 46: 2047-2052.

8. 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.

9. van de Weert-van Leeuwen PB, Slieker 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.

10. Ward JH. Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc* 1963; 58: 236-&.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

11. Godfrey S, Davies CT, Wozniak E, Barnes CA. Cardio-respiratory response to exercise in normal children. *Clin Sci* 1971; 40: 419-431.

12. Orenstein DM. Assessment of exercise pulmonary function. In: Rowland TW, editor. Pediatric Laboratory Exercise Testing Clinical Guidelines. Champaign/IL: Human Kinetics Publishers 1993. p. 141-163.

13. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J. 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.

14. 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.

15. 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.

16. MacKenzie T, Gifford AH, Sabadosa KA, Quinton HB, Knapp EA, Goss CH, Marshall BC. 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.

17. George PM, Banya W, Pareek N, Bilton D, Cullinan P, Hodson ME, Simmonds NJ. Improved survival at low lung function in cystic fibrosis: cohort study from 1990 to 2007. *BMJ* 2011; 342.

18. 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.

19. Radtke T, Hebestreit H, Gallati S, Schneiderman JE, Braun J, Stevens D, Hulzebos EHJ, Takken T, Boas SR, Urquhart DS, Lands LC, Tejero S, Sovtic A, Dwyer T, Petrovic M, Harris RA, Karila

C, Savi D, Usemann J, Mei-Zahav M, Hatziagorou E, Ratjen F, Kriemler S; 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.

20. McKone EF, Goss CH, Aitken ML. CFTR genotype as a predictor of prognosis in cystic fibrosis. *Chest* 2006; 130: 1441-1447.

21. de Gracia J, Mata F, Alvarez A, Casals T, Gatner S, Vendrell M, de la Rosa D, Guarner L, Hermosilla E. Genotype-phenotype correlation for pulmonary function in cystic fibrosis. *Thorax* 2005; 60: 558-563.

22. Weill D, Benden C, Corris PA, Dark JH, Davis RD, Keshavjee S, Lederer DJ, Mulligan MJ, Patterson GA, Singer LG, Snell GI, Verleden GM, Zamora MR, Glanville AR. 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.

23. Radtke T, Nevitt SJ, Hebestreit H, Kriemler S. Physical exercise training for cystic fibrosis. *Cochrane Database Syst Rev* 2017; 11: CD002768.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Figure 1. Flow chart of included patients.

Figure 2. Kaplan-Meier survival curve for three different VO_2peak groups.

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

Figure 4. Variable importance plot generated by the unsupervised Random Forest algorithm from the 12 clinical and physiolgocial parameters used for clustering. Higher values corresponed to higher importance of the variable for clustering. BMI, body mass index; CFRD, cystic fibrosis-related diabetes; CPET, cardiopulmonary exercise testing; FEV_1, forced expiratory volume in 1s; PA, Chronic *Pseudomonas aeruginosa* infection; PI, Pancreatic insufficiency; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV_1*35); VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; VO_2peak, peak oxygen consumption; Wpeak, peak work rate.

Table 1. Patients´ characteristics

	N=433
Mean Follow-up, all patients (years)	8.9±2.9 (0.1, 14.0)
Mean Follow-Up, survivors (years)	9.6±2.4 (5.0, 14.0)
Mean Follow-up, death or LTx (years)	5.9±3.3 (0.1, 13.5)
Death or LTx, N (%)	74 (17.1)
Age (years)	16.6±6.1 (10.0, 44.5)
Sex, N (% female)	184 (42.5)
BMI (z-score)	-0.70±1.0 (-4.53, 1.89)
FEV_1 (% predicted)	73.4±21.8 (19.7, 123.4)
Genotype	
CFTR, both allels from classes I-III, N (%)*	315 (72.7)
CFTR, at least one allele from classes IV-V, N (%)*	22 (5.1)
CFTR, at least one allele unknown/not available, N (%)	96 (22.2)
Comorbidities	
Chronic *Pseudomonas aeruginosa* infection, N (%)	295/424 (69.6)
Pancreatic insufficiency, N (%)	404/430 (93.3)
CFRD, N (%)	34/426 (8.0)
Exercise capacity	

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

VO_2peak (% predicted)	82.0±21.0 (23.3, 151.2)
Wpeak (% predicted)	91.2±23.2 (17.0, 197.0)

Data are mean±standard deviation (ranges) or number (%) of the study sample. BMI, body mass index; CFRD, cystic fibrosis-related diabetes; CFTR, cystic fibrosis transmembrane conductance regulator; LTx, lung transplantation; FEV_1, forced expiratory volume in 1s; VO_2peak, peak oxygen uptake; Wpeak, peak work rate. * Information on both CFTR mutations were available for 337 patients.

Table 2. Predictors of death or lung transplantation based on random-effects Cox proportional hazards regression (adjusted analyses – 10-year follow-up).

	CPET-derived variable in the model						
	VO_2peak (% pred)	Wpeak (% pred)	VE/VO_2peak	VE/VCO_2peak	VEpeak/MVVpred	SpO_2peak (%)	$\Delta VE/\Delta VCO_2$ 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 (years)	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_1 (% 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

Data are hazard ratios (95% confidence interval) and probability of a type I error. BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production relationship slope; SpO_2peak, oxygen saturation at peak exercise; VO_2peak, peak oxygen uptake; Wpeak, peak work rate.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Table 3 Cystic fibrosis risk groups defined by clinical and physiological parameters using Ward`s 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 (years)	15.91 ± 5.1$^{\$}$	14.5 ± 3.5$^{+,\#}$	17.1 ± 4.2$^{+,\P}$	29.9 ± 6.1$^{\$,\#,\P}$
BMI (z-score)	-0.17 ± 0.8$^{*\dagger,\$}$	-1.03 ± 0.8$^{*,+}$	-1.86 ± 1.1$^{\dagger,+,\P}$	-0.75 ± 0.8$^{\$,\P}$
Genotype**				
CFTR, both allels from classes I-III, N (%)	154 (74.4)	97 (74.6)	38 (70.4)	19 (57.6)
CFTR, at least one allele from classes IV-V, N (%)	14 (6.8)	6 (4.6)	1 (1.9)	0 (0)
CFTR, at least one allele unknown/not available, N (%)	39 (18.8)	27 (20.8)	15 (27.7)	14 (42.4)
CFRD	11 (5)$^{\$}$	10 (8)$^{\#}$	3 (6)¶	9 (27)$^{\$,\#,\P}$
Chronic PA colonization	120 (58)$^{*,\dagger,\$}$	97 (75)*	42 (78)†	30 (91)$^{\$}$
Pancreatic insufficiency	188 (91)$^{*,\dagger,\$}$	122 (94) *	53 (98)†	32 (97)$^{\$}$
Lung function				
FEV_1 (% predicted)	87.6 ± 13.2$^{*,\dagger,\$}$	69.2 ± 16.4$^{*,+,\#}$	43.1 ± 13.9$^{\dagger,+}$	49.3 ± 13.9$^{\$,\#}$
FVC (% predicted)	95.2 ± 11.1$^{*,\dagger,\$}$	80.9 ± 12.6$^{*,+,\#}$	60.9 ± 15.2$^{\dagger,+,\P}$	71.9 ± 15.6$^{\$,\#,\P}$
CPET variables				
VO_2peak (% predicted)	79.4 ± 16.2$^{*,\dagger,\$}$	64.8 ± 15.2$^{*,+}$	58.8 ± 15.3$^{\dagger,+,\P}$	73.4 ± 16.9$^{\$,\P}$
Wpeak (% predicted)	103.9 ± 20.1$^{*,\dagger,\$}$	83.2 ± 15.4$^{*,+}$	66.4 ± 21.3$^{\dagger,+,\P}$	84.3 ± 21.1$^{\$,\P}$
VE/VO_2peak	37.16 ± 5.6*,†	40.22 ± 6.5$^{*,+}$	48.01 ± 8.2$^{\dagger,+,\P}$	37.70 ± 6.9¶
VE/VCO_2peak	31.75 ± 4.5*,†	34.86 ± 5.3$^{*,+}$	41.58 ± 7.7$^{\dagger,+,\P}$	32.50 ± 4.2¶
SpO_2peak (%)	96.4 ± 3.1$^{\dagger,\$}$	95.3 ± 3.9$^{+,\#}$	91.5 ± 5.4$^{\dagger,+}$	91.1 ± 4.8$^{\$,\#}$
VEpeak/MVVpred	0.8 ± 0.2$^{*,\dagger,\$}$	0.9 ± 0.2$^{*,+}$	1.2 ± 0.2$^{\dagger,+,\P}$	1.1 ± 0.3$^{\$,\P}$
$\Delta VE/\Delta VCO_2$ slope	28.9 ± 4.1*,†	31.4 ± 4.3*	34.4 ± 8.6$^{\dagger,+}$	29.4 ± 3.7
Outcomes death or LTx				
Up to five years' follow-up	1 (0)$^{\dagger,\$}$	4 (3)$^{+}$	19 (35)$^{\dagger,+}$	4 (12)$^{\$}$
Up to ten years' follow-up	3 (1)$^{*,\dagger,\$}$	17 (13) $^{*,+}$	34 (63) $^{\dagger,+,\P}$	9 (27) $^{\$,\P}$
Entire study period	4 (2)$^{*,\dagger,\$}$	19 (15) $^{*,+,\#}$	36 (67)$^{\dagger,+}$	13 (39) $^{\$,\#}$

Data are mean±standard deviation or number (%). BMI, body mass index; CFRD, cystic fibrosis-related diabetes; CFTR, cystic fibrosis transmembrane conductance regulator; CPET, cardiopulmonary exercise test; FEV_1, forced expiratory volume in 1s; FVC, forced vital capacity; SpO_2peak, oxygen saturation at peak exercise; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV_1*35); VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; VO_2peak, peak oxygen consumption; Wpeak, peak work rate. Data for $\Delta VE/\Delta VCO_2$ slope was only available for a subset of patients (cluster 1=119, cluster 2=56, cluster 3=18, and cluster 4=21).

Differences among clusters were analysed using Chi-squared 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 6 as we compared 4 clusters). If the test passed the significance level, this is indicated by a sign [*],[†], [$], [+], [#], [¶]. ** Differences between clusters based on Chi-squared statistics (*P*=0.039). If Chi-square statistics is calculated for patients with known genotype by excluding patients with at least one "unknown; non-available" CFTR allele, no difference in genotype is observed between the groups (*P*=0.338).

[*] Difference between cluster 1 and 2.
[†] Difference between cluster 1 and 3.
[$] Difference between cluster 1 and 4.
[+] Difference between cluster 2 and 3.
[#] Difference between cluster 2 and 4.
[¶] Difference between cluster 3 and 4.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

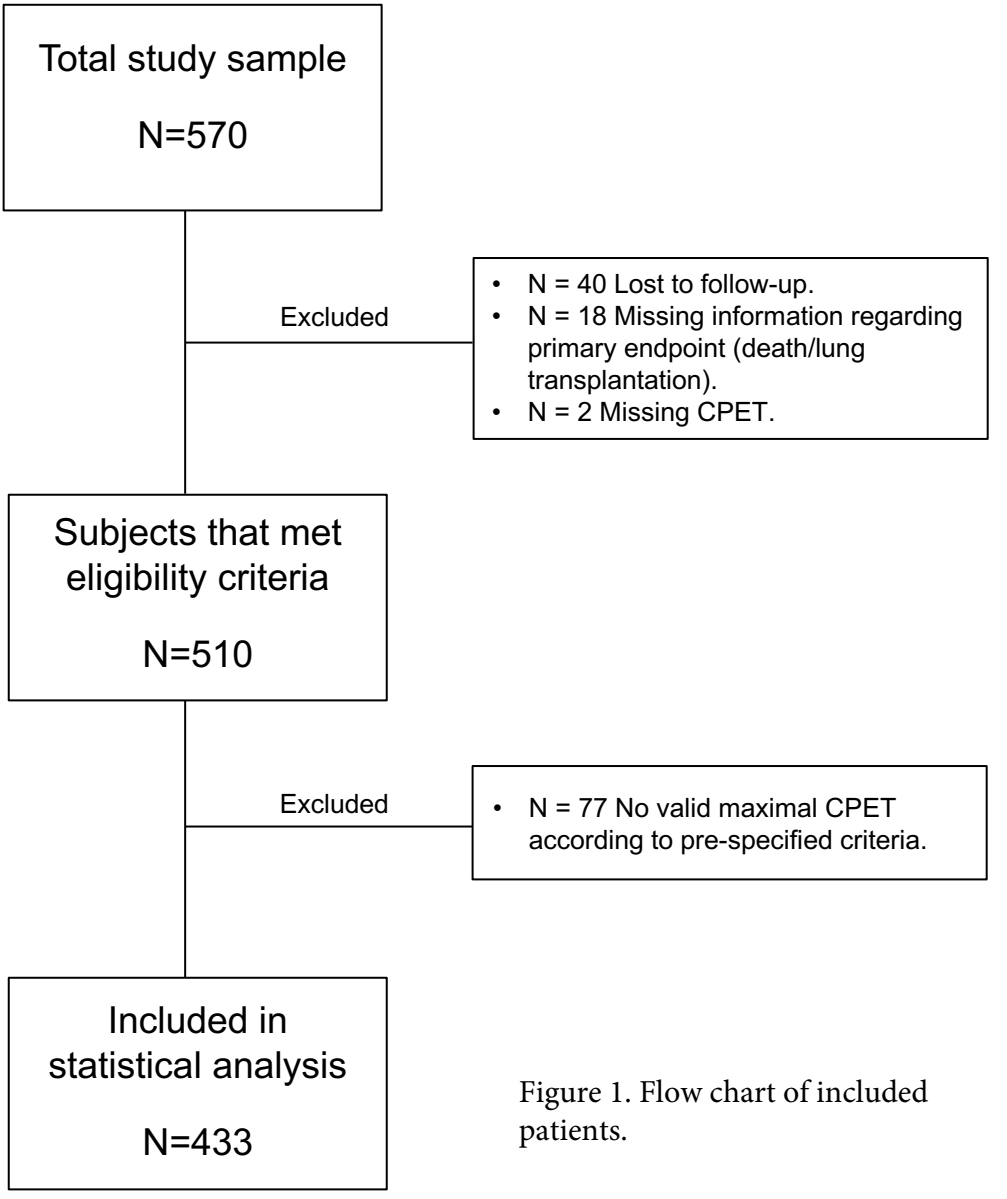

Figure 1. Flow chart of included patients.

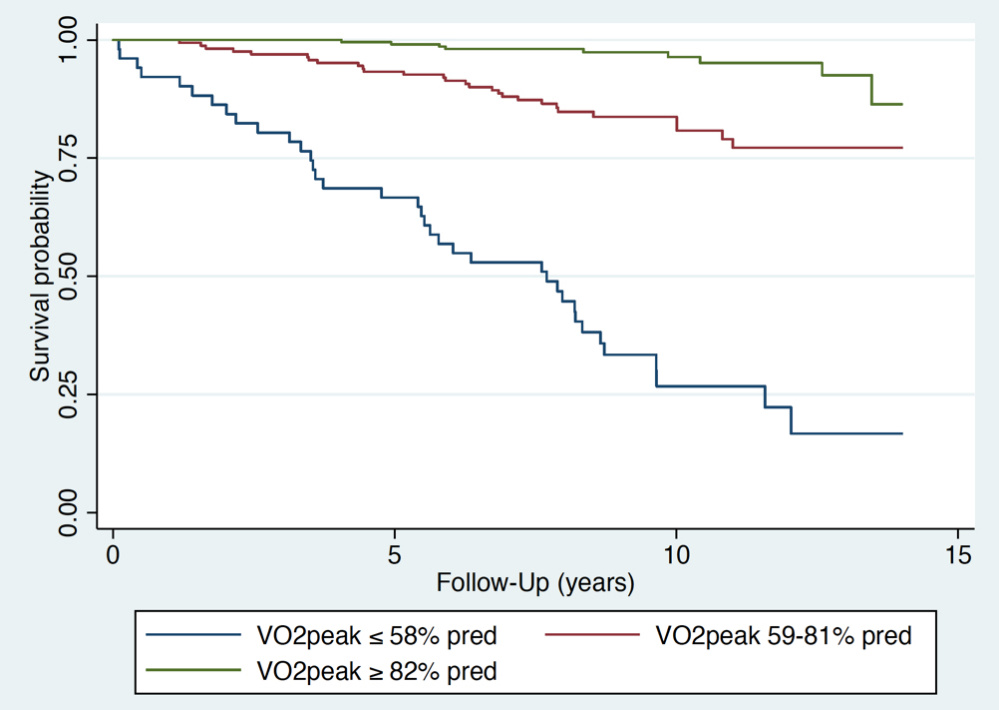

Figure 2. Kaplan-Meier survival curve for three different VO2peak groups.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

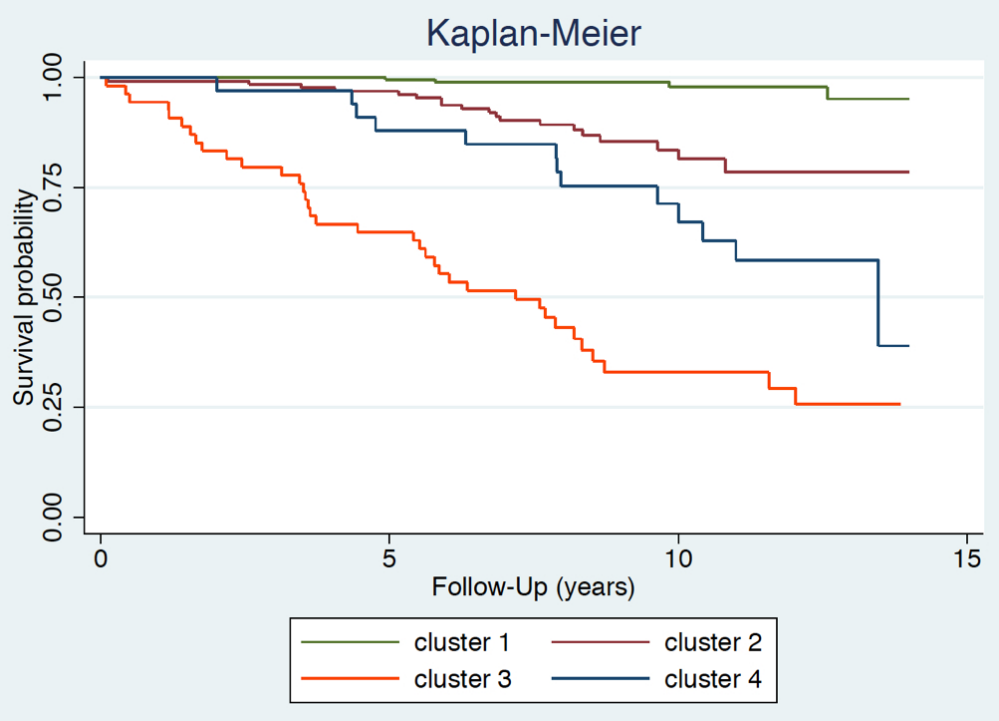

Figure 3. Kaplan-Meier survival curve for four different patient groups identified using Ward's hierarchical clustering.

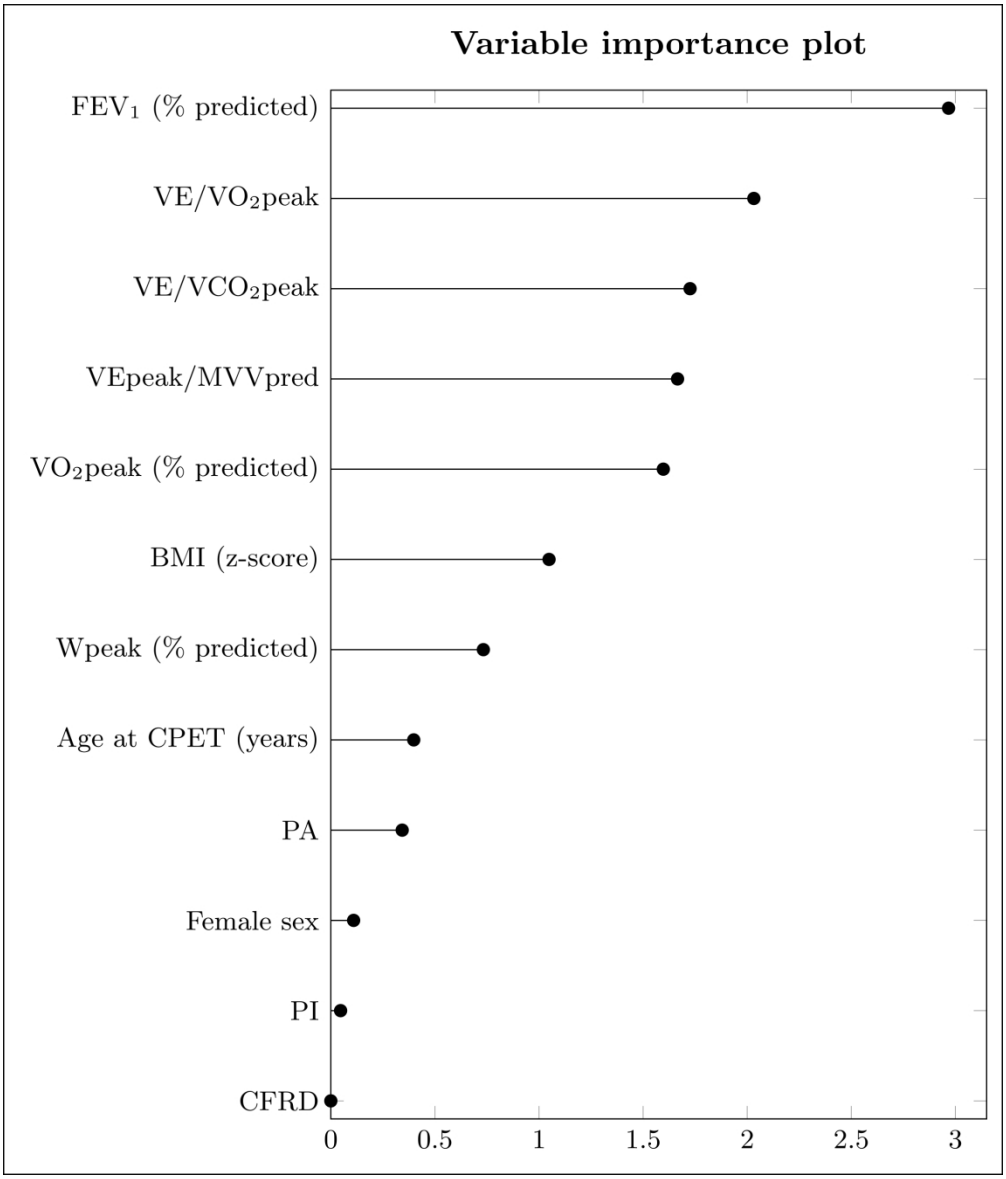

Figure 4. Variable importance plot generated by the unsupervised Random Forest algorithm from the 12 clinical and physiolgocial parameters used for clustering. Higher values corresponed to higher importance of the variable for clustering. BMI, body mass index; CFRD, cystic fibrosis-related diabetes; CPET, cardiopulmonary exercise testing; FEV1, forced expiratory volume in 1s; PA, Chronic Pseudomonas aeruginosa infection; PI, Pancreatic insufficiency; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV1*35); VE/VCO2peak, ventilatory equivalent for carbon dioxide; VE/VO2peak, ventilatory equivalent for oxygen; VO2peak, peak oxygen consumption; Wpeak, peak work rate.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

ONLINE DATA SUPPLEMENT

Cardiopulmonary exercise testing provides additional prognostic information in cystic fibrosis

Helge Hebestreit, Erik H.J. Hulzebos, Jane E. Schneiderman, Chantal Karila, Steven R. Boas, Susi Kriemler, Tiffany Dwyer, Margareta Sahlberg, Don S. Urquhart, Larry C. Lands, Felix Ratjen, Tim Takken, Liobou Varanistkaya, Viktoria Rücker, Alexandra Hebestreit, Jakob Usemann, Thomas Radtke for the Prognostic value of CPET in CF study group

Methods

Study design and subjects

For this retrospective study, we analyzed data of patients with CF aged ≥10 years who had a full CPET meeting prespecified criteria between 1st January 2000 and 31st December 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. We searched for publications on exercise testing in CF over the past 15 years to identify centres performing full CPET in CF using the Godfrey cycle ergometer protocol (1) or comparable (e.g., increase in exercise intensity minute-by-minute or as a ramp by 10-25 W/min) were identified. Publications on CPET in CF were searched in PubMed between November 2013 to February 2014 using the search terms "cardiopulmonary exercise testing" OR "exercise test" AND "cystic fibrosis". Sixteen possibly suitable centers were invited to participate in this study if they could provide data of at least 20 patients. Two centers never responded to repeated approaches, and four centers could not provide data fulfilling the above criteria.

Genotype classification

Cystic fibrosis transmembrane conductance regulator genotype (CFTR) was classified into five classes (27, 28). We grouped patients for descriptive analysis in those with both CFTR alleles in either class I, II or III (corresponding to severely reduced CFTR function) and patients with at least one mutant allele

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

in class IV or class V (corresponding to some residual CFTR function) (16), see Table 1 in the original publication. We further categorized patients into four CFTR groups: i) F508del homozygous, ii) F508del heterozygous, iii) others, and iv) unknown.

Ethical approval

Ethical approval was obtained from all respective ethical research committees, if required (see online supplements). Ethik-Kommission, Institut für Pharmakologie und Toxikologie (No 72/14), Würzburg, Wuerzburg, Germany; Medisch Ethische Toetsingscommissie (No 15-268/C), Utrecht, Netherlands; AMITA Health Institutional Review Board (#2015-006-03), Glenview, IL, USA; Sydney South West Area Health Service (RPAH Zone) Ethics Review Committee (No X14-0227), Sydney, Australia; Göteborgs Universitet, Medicinska fakultetens forkningsetikkommitté, Gothenborg, Sweden; Sick Kids Research Ethics Board, Toronto, Canada; Research and Development offices at University College London, London, UK; Mc Gill University Health Centre, Quebec, Montreal, Canada. For the centers in Paris and Switzerland, no ethical approval was required since the retrospective data provided were fully anonymized.

Clinical data

Centers were asked to provide the following data collected at the time of the CPET: age, sex, weight and height, pulmonary function, CPET-related data, genetic information, and binary coded CF-related comorbidities (exocrine pancreatic insufficiency, CFRD, colonization with *Pseudomonas aeruginosa*, *Burkholderia cepacia, Mycobacterium* other than tuberculosis), Chronic *Pseudomonas aeruginosa* infection was considered to be present when >50% of at least 4 sputum samples collected in the

previous year were positive (2) and the treating physician had had no reason to think that the status had changed.

Cardiopulmonary exercise testing

A valid exercise test with respect to this study was defined as a test with all required data available (see below), no equipment failure during the test, and indicators for a maximal effort. These included: i) the supervisors impression; ii) a rating of perceived exertion of 9 or 10 on a 0-10 Borg scale or at least 17 on a 6-20 Borg scale; iii) plateau in VO_2 (increase in VO_2 during the final completed stage of the test of less than 2 $mL \cdot kg^{-1} \cdot min^{-1}$); iv) a respiratory exchange ratio at peak exercise >1.03 in children or >1.05 in adults; v) a heart rate at peak exercise exceeding 195 $beats.min^{-1}$ in children or 209 – 0.86 × age in female adults or 207 – 0.78 × age in male adults; vi) VE at peak exercise exceeded estimated MVV (FEV_1 x 35), also known as breathing reserve index (3, 4).

The slope of the $\Delta VE/\Delta VCO_2$ -relationship below the respiratory compensation point was requested from centers that could access the source data in electronic format. Centers were asked to calculate the slope as suggested by Cooper at al. (5). Five centers provided data for the $\Delta VE/\Delta VCO_2$ slope.

An overview on equipment and exercise protocols used to perform the CPET is given in Table E1.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Table E1 Overview on equipment used to perform the cardiopulmonary exercise test (CPET).

Centre name	Equipment	Test protocol
Chicago/USA	**Ergometer:** Lode Excalibur; Groningen, the Netherlands **Metabolic Cart:** ZAN 680, nSpire, Louisville, Colorado; Quark B2 metabolic cart (Cosmed, Rome, Italy); Sensormedics 229 Breath by Breath System, Yorba Linda, California **Oxygen saturation:** Masimo RAD7, Irvine, CA, USA	Godfrey protocol (6)
Gothenburg	**Ergometer:** Electromagnetically braked cycle ergometer, Vmax 229 (Sensor Medics, Yorba Linda, CA, USA) **Metabolic Cart:** Oxygen tension and carbon dioxide tension were measured using a TCM™4 radiometer (Radiometer, Copenhagen, Denmark) **Oxygen saturation:** Pulse oximeter; MasimoSET Radical; Masimo Corp, Irvine, CA, USA)	Continuous increase in workload from 10-30 Watts per min depending on last year test.
London/UK	**Ergometer:** >128 and ≤135cm. = Ergoline 900 (Ergoline, Blitz, Germany);	Ramp protocol to achieve 8 to 12 minutes. incremental exercise test.

	>135cm. = Lode Excalibur (Lode, Groningen, Holland) **Metabolic cart:** MedGraphics (St Paul, Minnesota, USA). **Cardiac monitoring:** CardioControl Workstation (Welch Allyn, Delft, Holland). **Oxygen saturation:** Nonin (USA) SpO_2 probe placed over right supraorbital artery fixed with a bandanna	Expected workload of 3Watt/kg adjusted dependent upon the fitness quartile into which the child places themselves, and the pre-test spirometry values to attain a suitable ramping protocol
Montreal/Canada	**Ergometer:** Electronically-braked cycle ergometer (Sensormedics) **Metabolic cart:** Vmax, Cardinal Health **Oxygen saturation:** Reflectance probe (Massimo forehead probe) for pulse oximetry	Godfrey protocol (6)
Paris/France	**Ergometer:** Ergoline: Lode BV, Groningen, The Netherlands **Metabolic cart:** Gould 9000, Sensormedics, Dayton, DH ECG: Marquette Max-1, Marquette Electronics **Oxygen saturation:** Ohmeda 3700, Ohmeda, Louisville, CO	Individualized (from the predicted VO_2peak) incremental in exercise intensity (Watt) minute by minute
Switzerland	**Ergometer:** Ergoline 800c; (Pilger, St. Gallen, Switzerland) **Metabolic cart:** Quark B2 metabolic cart (Cosmed, Rome, Italy) **Oxygen saturation:** Nellcor Reflectance oxygen sensor RS10 & NPB290, Nellcor Pruitan Bennet Inc., Pleasanton, CA, USA)	Godfrey protocol (6) Stepwise increments depending on stature (<120 cm: 10 W/min; 120-150 cm: 15 W/min; >150 cm: 20 W/min; FEV_1<35%: 10W/min irrespective of stature)

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Sydney/Australia	**Ergometer:** Electronically-braked cycle ergometer (Ergometrics 800S, SensorMedics, USA); **Metabolic cart:** Breath-by-breath metabolic cart (VMax229, SensorMedics, USA); **Oxygen saturation:** Radical Massimo finger probe oximeter for pulse oximetery and pulse rate (Massimo Corporation, Irvine, USA)	Jones stage 1; stepwise increments depending on lung function, stature and reported exercise tolerance (5-30Watt/increment, aiming to reach peak exercise at 8-10 minutes)
Toronto/Canada	**Ergometer:** Lode Corival, Groningen, the Netherlands **Metabolic cart:** Vmax, Cardinal Health, **Oxygen saturation:** Reflectance probe (Massimo, forehead site) for pulse oximetry.	Godfrey Protocol (6)
Utrecht/Netherlands	**Ergometer:** Lode Corival, Groningen, the Netherlands **Metabolic cart:** Jaeger Oxycon (Germany) **Oxygen saturation:** Nellcor forehead sensor for pulse oximetry.	Godfrey Protocol (6)
Würzburg/Germany	**Ergometer:** Monark 834 E Ergomedic ergometer (Varberg, Sweden) **Metabolic cart:** CPX/D metabolic cart (MedGraphics, St Paul, MN) **Oxygen saturation:** Nelcor forehead sensor for pulse oximetry	Godfrey protocol (6) Stepwise increments depending on stature (<120 cm: 10 W/min; 120-150 cm: 15 W/min; >150 cm: 20 W/min; FEV_1<35%: 10W/min irrespective of stature)

FEV_1, forced expiratory volume in 1s; SpO_2, oxygen saturation; VO_2peak, peak oxygen uptake;

Statistical analyses

Cluster analysis methodology

An unsupervised, data-driven approach was used to explore the relevance of CPET-derived parameters to predict survival in CF patients. The following twelve parameters were investigated as input to perform Ward's hierarchical clustering with the software R version 3.0.2 (http://www.R-project.org) (7): forced expiratory volume in 1s (FEV_1), peak work rate (Wpeak), peak oxygen uptake (VO_2peak), age at CPET, the breathing reserve index at peak exercise (VEpeak/MVVpred), ventilatory equivalent for oxygen at peak exercise (VE/VO_2), ventilatory equivalent for carbon dioxide at peak exercise (VE/VCO_2), body mass index (BMI), sex, chronic *Pseudomonas aeruginosa* colonization, pancreatic insufficiency and cystic fibrosis-related diabetes (CFRD). The separation between clusters was calculated using the Euclidean distance, and the agglomeration procedure was done according to Ward's minimum variance method. The optimal number of clusters was determined based on the majority rule for indices using the R-package "NbClust" (8). In the next step, the relevance of each parameter used for clustering was investigated. Therefore, an unsupervised Random Forests analysis approach was chosen where the data is classified without a priori classification specifications (9). For each variable, an importance value is reported. The variables are ranked within a variable importance plot according to their importance.

Based on the Forest plot, the parameters for clustering were selected. Given that some of the input parameters were correlated, principal component analysis (PCA) was performed as a dimension reduction procedure. The resulting linear combination corresponds to a principal component (PC) (10). The number of PC were selected that would explain at least 95% of overall variation in data. Finally, these components were used as input parameters for Ward's hierarchical clustering, as described above.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Results

Results of Ward's hierarchical clustering

All CPET parameters and FEV_1 had a high variable importance to define clusters, as indicated by the Forest plot (Figure 4 in the original publication). However, the binary coded variables (sex, chronic *Pseudomonas aeruginosa* colonization, pancreatic insufficiency and CFRD were less important. Based on the Forest plot, the eight continuous variables (FEV_1, BMI, Wpeak, age at CPET, VO_2peak, VEpeak/MVVpred, VE/VO_2peak, and VE/VCO_2peak) were introduced into the PCA, and five orthogonal factors (=principal components) explained more than 95% of variance (Table E12). Ward's hierarchical clustering was feasible, and the optimal number of clusters was four. The heat map containing the dendogram obtained using clustering of the five PCs is shown in Figure E2. The characteristics of the four clusters identified containing n=207, n=130, n=54 and n=33 patients is given in Table 3 in the original publication.

Table E2 Comparison of genotype, lung function, comorbidities and cardiopulmonary exercise testing data among survivors and cases

Variables	**Survivors** (N=359)	**Death/LTx** (N=74)	***P*-value**
Age (years)	14.2 (12.4, 17.7)	17.1 (14.0, 22.5)	<0.001
Sex, N female (%)	146/359 (41)	38/74 (50)	0.091
BMI (z-score)	-0.55±0.90	-1.39±1.14	<0.001
Genotype			
CFTR, both alleles from classes I-III, N (%)	271/359 (76)	44/74 (59)	<0.001[#]
CFTR, at least one allele from classes IV-V, N (%)	21/359 (6)	1/74 (1)	
CFTR, at least one allele unknown/not available, N (%)	67/359 (19)	29/74 (39)	
F508del homozygous, N (%)	167/359 (47)	34/74 (46)	0.003[$]
F508del heterozygous, N (%)	137/359 (38)	19/74 (26)	
Others, N (%)	36/359 (10)	9/74 (12)	
Unknown, N (%)	19/359 (5)	12/74 (16)	
Lung function			
FEV_1 (% predicted)	80.5 (67.3, 91.4)	45.6 (34.2, 55.5)	<0.001
FVC (% predicted)	88.7±14.6	65.0±16.3	<0.001

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Comorbidities			
Chronic *Pseudomonas aeruginosa* infection (%)	233/356 (65)	62/68 (91)	<0.001
Pancreatic insufficiency (%)	331/357 (93)	74/74 (100)	0.017
CFRD (%)	24/357 (7)	10/69 (14)	0.029
CPET parameters			
VO_2peak (% predicted)	86.7±18.3	59.2±19.1	<0.001
Wpeak (% predicted)	94.1 (80.9, 107.9)	66.3 (53.0, 81.1)	<0.001
VE/VO_2peak	38.0 (34.6, 43.4)	40.2 (35.6, 48.5)	0.013
VE/VCO_2peak	33.0 (29.6, 36.9)	35.7 (31.2, 42.1)	0.001
VEpeak/MVVpred	0.88 (0.76, 1.06)	1.06 (0.94, 1.20)	<0.001
SpO_2peak (%)	96.0 (94.0, 98.0)	93.0 (88.0, 95.0)	<0.001
$\Delta VE/\Delta VCO_2$ slope*	29.5 (26.5, 32.9)	30.4 (27.5, 34.8)	0.193

Data are mean±SD, median (interquartile range) or N (%). BMI, body mass index; CFRD, cystic fibrosis-related diabetes; CFTR, cystic fibrosis transmembrane conductance regulator; FEV_1, forced expiratory volume in 1s; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production relationship slope; SpO_2peak, oxygen saturation at peak exercise; VO_2peak, peak oxygen uptake; Wpeak, peak work rate. * Data only available for 218 patients (24 death/LTx cases). Comparisons between groups were done using the independent t-test; Mann-Whitney-U test or the Chi-square test, as appropriate. [#] If Chi-square statistics is calculated for patients with known genotype by excluding patients with at least one "unknown; non-available" CFTR allele, no difference in genotype is observed between the groups (P=0.21). [$] If Chi-square statistics is calculated for patients with known genotype by excluding the "unknown", no difference in genotype is observed between the groups (P=0.33).

Table E3. Patients´ characteristics separated by study center.

Center	N	Mean Follow-up (years)	Death / LTx N (%)	Age (years)	FEV_1 (% predicted)	BMI (z-score)	Chronic PA (%)	PI (%)	CFRD (%)
Chicago	44	9.5±1.7	3 (7)	16.3±3.6	78.4±20.1	-0.45±1.19	79.5	93.2	2.3
Gothenburg	24	10.9±1.6	0 (0)	25.6±5.9	84.7±21.6	-0.36±0.83	54.2	79.2	20.8
London	24	6.8±2.3	6 (25)	13.4±1.8	70.4±21.3	-0.54±0.68	95.8	91.7	4.2
Montreal	18	7.4±0.9	1 (6)	13.2±1.4	81.0±19.3	-0.58±0.65	38.9	100.0	11.1
Paris	44	7.8±3.4	14 (32)	15.0±3.0	68.4±24.2	-0.90±0.99	94.3	92.7	10.8
Swiss	37	12.3±3.0	7 (19)	21.8±5.4	61.4±20.4	-1.51±1.04	89.2	97.3	13.5
Sydney	28	7.4±3.2	15 (54)	27.3±6.9	53.1±22.0	-0.90±0.83	92.9	92.9	25.0
Toronto	72	9.7±2.1	4 (6)	12.1±1.7	83.8±15.4	-0.35±0.94	18.1	88.9	6.9
Utrecht	74	7.0±1.4	7 (9)	13.3±1.1	73.3±17.6	-0.81±0.80	95.9	100.0	2.7
Würzburg	68	9.9±2.9	17 (25)	18.1±6.1	72.9±23.5	-0.66±1.11	61.2	97.1	3.0
Total	**433**	**8.9±2.9 (0.1; 14.0)**	**74 (17)**	**16.6±6.1 (10.0; 44.5)**	**73.4±21.8 (1.7; 123.4)**	**-0.70±1.0 (-4.53; 1.89)**	**69.6**	**94.0**	**8.0**

Data are mean±standard deviation or number (percentage) or ranges of study sample. BMI, body mass index; CFRD, cystic fibrosis-related diabetes; LTx, lung transplantation; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; PI, Pancreatic insufficiency. Death/LTx rates were significantly different between centers (Chi-squared test, $P<0.001$).

Table E4 Cardiopulmonary exercise testing variables separated by study center.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Center	N	VO_2peak (% predicted)	Wpeak (% predicted)	VE/VO_2peak	VE/VCO_2peak	VEpeak/MVVpred (%)	SpO_2peak (%)	$\Delta VE/\Delta VCO_2$slope
Chicago	44	100.0±25.3	96.8±22.7	41.3±7.2	35.3±5.7	98.1±23.7	93.6±2.6	33.8±7.6
Gothenburg	24	91.4±20.1	117.0±24.8	41.3±5.8	31.8±3.4	93.6± 2.0	95.1±2.0	27.2±2.8
London	24	76.7±15.3	91.5±20.0	39.1±7.1	33.3±6.6	95.0±24.8	96.6±3.6	---
Montreal	18	94.5±20.8	85.2±13.5	35.0±3.8	29.6±1.7	86.8±24.7	96.2±2.4	---
Paris	44	68.5±19.4	73.0±18.8	46.5±7.7	40.6±7.7	101.6±32.7	93.6±3.2	---
Swiss	37	77.4±17.1	102.7±20.9	42.4±6.7	37.4±5.2	110.5±26.5	89.5±4.6	31.4±7.0
Sydney	28	54.1±18.8	64.9±24.8	37.2±8.5	31.6±6.8	80.1±21.2	92.7±4.9	---
Toronto	72	90.0±14.2	84.7±13.1	40.9±4.9	35.8±3.6	98.2±21.6	98.9±1.3	30.3±3.5
Utrecht	74	81.8±15.0	98.3±20.0	36.8±6.9	31.7±5.4	91,5±19.4	96.0±3.0	---
Würzburg	68	79.7±17.8	95.6±21.9	35.1±5.7	30.7±4.8	83.4±22.9	94.7±5.0	27.2±3.5
Total	**433**	**82.0±21.1**	**91.2±23.3**	**39.5±7.3**	**34.0±6.2**	**94.0±25.3**	**95.1±4.3**	**30.2±5.7**
		(23.3, 151.2)	**(17.3, 196.8)**	**(13.3, 69.2)**	**(11.1, 68.4)**	**(23.5, 206.1)**	**(74, 100)**	**(19.7, 72.0)**

Data are mean±standard deviation or number (percentage) or ranges of study sample. $\Delta VE/\Delta VCO_2$slope, minute ventilation-carbon dioxide production relationship slope; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV_1*35); VE/VCO_2, ventilatory equivalent for carbon dioxide; VE/VO_2, ventilatory equivalent for oxygen; SpO_2peak, oxygen saturation at peak exercise; VO_2peak, peak oxygen uptake; Wpeak, peak work rate.

Table E5. Cox proportional hazards regression predictors of death or lung transplantation (LTx). Univariate analyses using a random effects model with study center adjustment.

Variable	Total N	Hazard ratio (exp(B))	*P*-value	Interpretation
Female Sex	433	1.490 (0.937; 2.370)	0.092	Female sex (n=184) not significantly worse outcome than male sex (n=249)
CFRD	426	1.649 (0.819; 3.319)	0.161	CFRD (n=34) not significantly worse outcome than no CFRD (n=392)
Chronic *Pseudomonas aeruginosa* infection	424	4.697 (1.889; 11.678)	**0.001**	Chronic PA (n=295) significantly worse outcome than no PA (n=129)
Age at CPET (years)	433	1.009 (0.969; 1.051)	0.661	Higher age at CPET not significantly associated with unfavorable or favorable outcome
FEV_1 (% predicted)	431	0.937 (0.924; 0.950)	**0.001**	Higher FEV_1 significantly associated with better outcome
BMI (z-score)	433	0.534 (0.438; 0.651)	**<0.001**	Higher BMI significantly associated with better outcome
VO_2peak (% predicted)	433	0.935 (0.924; 0.947)	**<0.001**	Higher VO_2peak significantly associated with better outcome
Wpeak (% predicted)	426	0.944 (0.932; 0.956)	**<0.001**	Higher Wpeak significantly associated with better outcome
VE/VO_2peak	433	1.084 (1.050; 1.119)	**<0.001**	Higher VE/VO_2 significantly associated with unfavorable outcome
VE/VCO_2peak	433	1.125 (1.085; 1.167)	**<0.001**	Higher VE/VCO_2 significantly associated with unfavorable outcome
VEpeak/MVVpred	432	12.656 (5.825; 27.498)	**<0.001**	Higher VEpeak/MVVpred significantly associated with unfavorable outcome
SpO_2peak (%)	393	0.866 (0.830; 0.904)	**<0.001**	Higher SpO_2peak significantly associated with better outcome
$\Delta VE/\Delta VCO_2$ slope	218	1.128 (1.072; 1.186)	**<0.001**	Higher $\Delta VE/\Delta VCO_2$ slope significantly associated with unfavorable outcome

BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production relationship slope; SpO_2peak, oxygen saturation at peak exercise; VO_2peak, peak oxygen uptake; Wpeak, peak work rate. Sex is coded as 0=male and 1=female. Effect estimates for pancreatic insufficiency could not be computed because there were no cases (death/LTx) that were pancreatic sufficient.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Table E6. Cox proportional hazards regression predictors of death or lung transplantation (LTx). Univariate analyses using a random effects model without study center adjustment.

Variable	Total N	Hazard ratio (95% CI)	*P*-value	Interpretation
Female Sex	433	1.481 (0.939; 2.336)	0.092	Female sex (n=184) not significantly worse than male sex (n=249)
CFRD	426	1.922 (0.983; 3.759)	0.056	CFRD (n=34) not significantly worse than no CFRD (n=392)
Chronic *Pseudomonas aeruginosa*	424	5.263 (2.275; 12.173)	**<0.001**	Chronic PA (n=295) significantly worse than no PA (n=129)
Age at CPET (years)	433	1.031 (1.000; 1.064)	0.053	Higher age at CPET not significantly associated with unfavorable or favorable outcome
FEV_1 (% predicted)	431	0.936 (0.924; 0.948)	**0.001**	Higher FEV_1 significantly associated with better outcome
BMI (z-score)	433	0.546 (0.454; 0.655)	**<0.001**	Higher BMI significantly associated with better outcome
VO_2peak (%predicted)	433	0.935 (0.924; 0.947)	**<0.001**	Higher VO_2peak significantly associated with better outcome
Wpeak (%predicted)	426	0.941 (0.930; 0.953)	**<0.001**	Higher Wpeak significantly associated with better outcome
VE/VO_2peak	433	1.059 (1.028; 1.092)	**<0.001**	Higher VE/VO_2 significantly associated with unfavorable outcome
VE/VCO_2peak	433	1.095 (1.058; 1.134)	**<0.001**	Higher VE/VCO_2 significantly associated with unfavorable outcome
VEpeak/MVVpred	432	6.188 (2.994; 12.791)	**<0.001**	Higher VEpeak/MVVpred significantly associated with unfavorable outcome
SpO_2peak (%)	393	0.878 (0.845; 0.912)	**<0.001**	Higher SpO_2peak significantly associated with better outcome
$\Delta VE/\Delta VCO_2$ slope	218	1.108 (1.055-1.164)	**<0.001**	Higher $\Delta VE/\Delta VCO_2$ slope significantly associated with unfavorable outcome

BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; SpO_2peak, oxygen saturation at peak exercise; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV_1*35); VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation -carbon dioxide production relationship slope; VO_2peak, peak oxygen uptake; Wpeak, peak work rate. Sex is coded as 0=male and 1=female.

Table E7. Cox proportional hazards regression predictors of death or lung transplantation (LTx). Univariable analyses using a random effects model with study center adjustment in a subsample of patients with advanced lung disease ($FEV_1 \leq 40\%$ predicted).

Variable	Total N	Hazard ratio (exp(B))	*P*-value	Interpretation
Female Sex	39	1.023 (0.408; 2.564)	0.961	Female sex not significantly worse outcome than male sex
CFRD	36	1.079 (0.364; 3.120)	0.891	CFRD not significantly worse outcome than no CFRD
Chronic *Pseudomonas aeruginosa* infection	37	1.760 (0.232; 13.350)	0.584	Chronic PA not significantly worse outcome than no PA
Age at CPET (years)	39	0.910 (0.844; 0.981)	**0.014**	Higher age at CPET significantly associated with favorable outcome
FEV_1 (% predicted)	39	0.964 (0.891; 1.043)	0.364	Higher FEV_1 not significantly associated with better outcome
BMI (z-score)	39	0.923 (0.669; 1.273)	0.624	Higher BMI not significantly associated with better outcome
VO_2peak (% predicted)	39	0.962 (0.931; 0.994)	**0.019**	Higher VO_2peak significantly associated with better outcome
Wpeak (% predicted)	39	0.963 (0.936; 0.990)	**0.007**	Higher Wpeak significantly associated with better outcome
VE/VO_2peak	39	1.081 (1.03; 1.114)	**0.002**	Higher VE/VO_2 at peak exercise significantly associated with unfavorable outcome
VE/VCO_2peak	39	1.092 (1.033; 1.154)	**0.002**	Higher VE/VCO_2 at peak exercise significantly associated with unfavorable outcome
VEpeak/MVVpred	39	0.556 (0.141; 2.191)	0.401	Higher VEpeak/MVVpred not significantly associated with unfavorable outcome
SpO_2peak (%)	32	0.939 (0.863; 1.022)	0.147	Higher SpO_2peak not significantly associated with better outcome
$\Delta VE/\Delta VCO_2$ slope	14	1.042 (0.977; 1.113)	0.208	Higher $\Delta VE/\Delta VCO_2$ slope not significantly associated with unfavorable outcome

BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production

Table E8. Cox proportional hazards regression predictors of death or lung transplantation (LTx). Univariable analyses using a random effects model without study center adjustment in a subsample of patients with advanced lung disease ($FEV_1 \leq 40\%$ predicted).

Variable	Total N	Hazard ratio (exp(B))	*P*-value	Interpretation

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Female Sex	39	0.894 (0.385; 2.076)	0.794	Female sex not significantly worse outcome than male sex
CFRD	36	1.079 (0.364; 3.199)	0.891	CFRD not significantly worse outcome than no CFRD
Chronic *Pseudomonas aeruginosa* infection	37	1.833 (0.246; 13.637)	0.554	Chronic PA not significantly worse outcome than no PA
Age at CPET (years)	39	0.954 (0.900; 1.011)	0.114	Higher age at CPET not significantly associated with favorable or unfavorable outcome
FEV_1 (% predicted)	39	0.952 (0.881; 1.031)	0.228	Higher FEV_1 not significantly associated with better outcome
BMI (z-score)	39	0.982 (0.737; 1.309)	0.902	Higher BMI not significantly associated with better outcome
VO_2peak (% predicted)	39	0.961 (0.931; 0.992)	**0.014**	Higher VO_2peak significantly associated with better outcome
Wpeak (% predicted)	38	0.963 (0.938; 0.989)	**0.006**	Higher Wpeak significantly associated with better outcome
VE/VO_2peak	39	1.082 (1.030; 1.137)	**0.002**	Higher VE/VO_2 at peak exercise significantly associated with unfavorable outcome
VE/VCO_2peak	39	1.092 (1.035; 1.153)	**0.001**	Higher VE/VCO_2 at peak exercise significantly associated with unfavorable outcome
VEpeak/MVVpred	39	0.860 (0.265; 2.791)	0.801	Higher VEpeak/MVVpred not significantly associated with unfavorable outcome
SpO_2peak (%)	32	0.966 (0.899; 1.038)	0.350	Higher SpO_2peak not significantly associated with better outcome
$\Delta VE/\Delta VCO_2$ slope	14	1.043 (0.977; 1.112)	0.208	Higher $\Delta VE/\Delta VCO_2$ slope not significantly associated with unfavorable outcome

BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production

Table E9. Cox proportional hazards regression predictors of death or lung transplantation (LTx). Univariable analyses using a random effects model with study center adjustment restricted to 2-year follow-up data.

Variable	Total N	Hazard ratio (exp(B))	*P*-value	Interpretation
Female Sex	433	1.331 (0.380; 4.658)	0.654	Female sex not significantly worse outcome than male sex

CFRD	426	3.120 (0.625; 16.395)	0.163	CFRD not significantly worse outcome than no CFRD
Chronic *Pseudomonas aeruginosa* infection	424	-	-	All death/LTx cases were chronically infected with *Pseudomonas* aeruginosa. Hazard ratio could therefore not be calculated.
Age at CPET (years)	433	1.008 (0.900; 1.128)	0.895	Higher age at CPET not significantly associated with favorable or unfavorable outcome
FEV_1 (% predicted)	431	0.916 (0.877; 0.957)	**<0.001**	Higher FEV_1 significantly associated with better outcome
BMI (z-score)	433	0.380 (0.218; 0.664)	**0.001**	Higher BMI significantly associated with better outcome
VO_2peak (% predicted)	433	0.921 (0.889; 0.955)	**<0.001**	Higher VO_2peak significantly associated with better outcome
Wpeak (% predicted)	426	0.934 (0.908; 0.961)	**<0.001**	Higher Wpeak significantly associated with better outcome
VE/VO_2peak	433	1.151 (1.075; 1.231)	**<0.001**	Higher VE/VO_2 at peak exercise significantly associated with unfavorable outcome
VE/VCO_2peak	433	1.170 (1.094; 1.250)	**<0.001**	Higher VE/VCO_2 at peak exercise significantly associated with unfavorable outcome
VEpeak/MVVpred	432	16.780 (3.092; 91.064)	**0.001**	Higher VEpeak/MVVpred significantly associated with unfavorable outcome
SpO_2peak (%)	393	0.822 (0.733; 0.921)	**0.001**	Higher SpO_2peak significantly associated with better outcome
$\Delta VE/\Delta VCO_2$ slope	218	0.768 (0.437; 1.348)	0.358	Higher $\Delta VE/\Delta VCO_2$ slope not significantly associated with unfavorable outcome

BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production

Table E10. Cox proportional hazards regression predictors of death or lung transplantation (LTx). Univariable analyses using a random effects model without study center adjustment restricted to 2-year follow-up data.

Variable	Total N	Hazard ratio (exp(B))	*P*-value	Interpretation
Female Sex	433	1.367 (0.396; 4.724)	0.621	Female sex not significantly worse outcome than male sex
CFRD	426	3.871 (0.781; 19.179)	0.097	CFRD not significantly worse outcome than no

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

Chronic *Pseudomonas aeruginosa* infection	424	-	-	CFRD All death/LTx cases were chronically infected with *Pseudomonas* aeruginosa. Hazard ratio could therefore not be calculated.
Age at CPET (years)	433	1.021 (0.932; 1.118)	0.657	Higher age at CPET not significantly associated with favorable or unfavorable outcome
FEV_1 (% predicted)	431	0.914 (0.875; 0.955)	**<0.001**	Higher FEV_1 significantly associated with better outcome
BMI (z-score)	433	0.440 (0.280; 0.691)	**<0.001**	Higher BMI significantly associated with better outcome
VO_2peak (% predicted)	433	0.921 (0.889; 0.955)	**<0.001**	Higher VO_2peak significantly associated with better outcome
Wpeak (% predicted)	426	0.936 (0.913; 0.960)	**<0.001**	Higher Wpeak significantly associated with better outcome
VE/VO_2peak	433	1.160 (1.084; 1.240)	**<0.001**	Higher VE/VO_2 at peak exercise significantly associated with unfavorable outcome
VE/VCO_2peak	433	1.170 (1.094; 1.250)	**<0.001**	Higher VE/VCO_2 at peak exercise significantly associated with unfavorable outcome
VEpeak/MVVpred	432	23.22 (4.565; 118.127)	**<0.001**	Higher VEpeak/MVVpred significantly associated with unfavorable outcome
SpO_2peak (%)	393	0.852 (0.778; 0.934)	**0.001**	Higher SpO_2peak significantly associated with better outcome
$\Delta VE/\Delta VCO_2$ slope	218	0.768 (0.437; 1.348)	0.358	Higher $\Delta VE/\Delta VCO_2$ slope not significantly associated with unfavorable outcome

BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; PA, *Pseudomonas aeruginosa*; VEpeak/MVVpred, breathing reserve index; VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; $\Delta VE/\Delta VCO_2$ slope, minute ventilation-carbon dioxide production

Table E11 Cystic fibrosis risk groups based on three different VO_2peak groups according to Nixon et al. (11)

	VO_2peak ≥82 predicted N=239 (group 1)	**VO_2peak 59-81 predicted N=133 (group 2)**	**VO_2peak ≤58 predicted N=61 (group 3)**	***P*-value**
Exposure Variable				
Female Sex	98 (41)	57 (43)	29 (48)	0.357
Age at CPET (years)	15.7 ± 5.4	16.8 v 6.0$	20.0 ± 7.6	< 0.001
BMI (z-score)	-0.42 ± 0.83	-0.80 ± 0.93$	-1.59 ± 1.17	< 0.001
CFRD	13 (5)†	13 (10)	8 (13)	0.052

Chronic *Pseudomonas a.* colonization	137 (57)†	106 (80)$	52 (85)	< 0.001
FEV_1 (% predicted)	84.4 ±15.0*†	65.8 ± 19.7$	47.1 ± 19.4	< 0.001
FVC (% predicted)	92.9 ± 11.9*†	79.3 ± 16.0$	63.8 ± 16.2	< 0.001
VO_2peak (% predicted)	96.8 ±13.3*†	71.1 ± 5.9$	47.5 ± 9.8	< 0.001
Wpeak (% predicted)	103.8 ± 19.1*†	82.1 ± 13.1$	60.2 ± 17.9	< 0.001
VE/VO_2peak	38.2 ± 6.2*†	39.9 ± 6.9	43.3 ± 10.4	0.025
VE/VCO_2peak	32.7 ± 4.9†	34.7 ± 5.4$	37.5 ± 9.5	0.004
SpO_2peak (%)	96.2 ± 3.5*†	94.5 ± 4.1	91.5 ± 5.7	< 0.001
VEpeak/MVVpred	0.92 ± 0.20	0.96 ± 0.30$	0.96 ± 0.32	0.601
$\Delta VE/\Delta VCO_2$ slope	29.4 ± 4.4	30.5 ± 4.6	36.0 ± 13.9	
Outcomes death or LTx				
Up to five years' follow-up	2 (1)*†	10 (8)$	18 (30)	< 0.001
Up to ten years' follow-up	6 (3)*†	22 (17)$	37 (61)	< 0.001
Entire study period	10 (4)*†	25(19)$	39 (65)	< 0.001

Values are N (percent) or mean±standard deviation. Differences in the distribution of characteristics across phenotypes were assessed using Chi-squared tests for categorical variables, and Kruskal–Wallis for continuous variables. BMI, body mass index; CFRD, cystic fibrosis-related diabetes; FEV_1, forced expiratory volume in 1s; FVC, forced vital capacity; SpO_2peak, oxygen saturation at peak exercise; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV_1*35); VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; VO_2peak, peak oxygen uptake; Wpeak, peak work rate. *Difference between group 1 and 2; †Difference between group 1 and 3; $Difference between group 2 and 3 using Chi-squared tests for categorical variables and Kruskal–Wallis, as appropriate. The Bonferroni-corrected significance level for these tests was 0.017 (overall significance level (0.05) divided by number of tests, which was 3 as we compared 4 clusters). If the test passed the significance level, a sign is shown in the table. Data for $\Delta VE/\Delta VCO_2$ slope was only available for a subset of patients (group 1=143; group 2=60 and group 3=15.

Table E12 Correlation of the original continues variables with the five main components derived from the prinicpal component analysis

	PC 1	**PC 2**	**PC 3**	**PC 4**	**PC 5**
Age at CPET (years)	-0,1536	-0,4674	0,4061	0,7594	-0,0196
FEV_1 (% predicted)	0,7597	0,3782	-0,3885	0,2027	-0,1916
BMI (z-score)	0,6162	0,1988	-0,0788	0,2119	0,7227
VO_2peak (% predicted)	0,7526	0,4606	0,3293	-0,1013	-0,0568
Wpeak (% predicted)	0,6776	0,3321	0,4841	0,1289	-0,2676
VE/VO_2	-0,6188	0,6752	-0,1291	0,3127	-0,0608

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

VE/VCO_2	-0,6852	0,6392	-0,1640	0,1904	0,0060
VEpeak/MVVpred	-0,5551	0,3556	0,6522	-0,2612	0,2022

CPET, cardiopulmonary exercise test; FEV_1, forced expiratory volume in 1s; BMI, body mass index; PC, Principal Component; VEpeak/MVVpred, breathing reserve index (MVV was calculated as FEV1*35); VE/VCO_2peak, ventilatory equivalent for carbon dioxide; VE/VO_2peak, ventilatory equivalent for oxygen; VO_2peak, peak oxygen uptake; Wpeak, peak work rate. Five orthogonal factors extracted explained >95% of variance.

Figure legends

Figure E1. Kaplan-Meier survival curve for three different Wpeak groups. Wpeak, peak work rate.

Figure E2. Heat map representing hierarchical Ward's clustering. The left color bar denotes individual subject grouping and their related cluster indicated by different colors: cluster one (Cyan, n=207), cluster two (green, n=130), cluster three (red n=54) and cluster four (purple, n=33). Column on the right: colors correspond to the clusters reported on the left bar, and for each cluster the percent of the main outcome (deaths or LTx) is reported. Horizontal bar denotes the values from each principal component (PC): dark blue represents values below the mean, white represents the mean value, and dark red represents values above the mean.

References

E1. Godfrey S, Mearns M. Pulmonary function and response to exercise in cystic fibrosis. *Arch Dis Child* 1971; 46: 144-151.

E2. Ballmann M, Rabsch P, von der Hardt H. Long-term follow up of changes in FEV1 and treatment intensity during Pseudomonas aeruginosa colonisation in patients with cystic fibrosis. *Thorax* 1998; 53: 732-737.

E3. ATS/ACCP. Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003; 167: 211-277.

E4. Hebestreit H, Arets HG, Aurora P, Boas S, Cerny F, Hulzebos EH, Karila C, Lands LC, Lowman JD, Swisher A, Urquhart DS, European Cystic Fibrosis Exercise Working G. Statement on Exercise Testing in Cystic Fibrosis. *Respiration* 2015; 90: 332-351.

E5. Cooper DM, Kaplan MR, Baumgarten L, Weiler-Ravell D, Whipp BJ, Wasserman K. Coupling of ventilation and CO2 production during exercise in children. *Pediatr Res* 1987; 21: 568-572.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

E6. Godfrey S. Exercise tests in assessing children with lung or heart disease. *Thorax* 1970; 25: 258.

E7. Team RC. R: A language and environment for statistical computing. R Foundation for Statistcial Computing. Vienna, Austria; 2017.

E8. Charrad M, Ghazzali N, Boiteau V, Niknafs A. Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *J Stat Softw* 2014; 61: 1-36.

E9. Breiman L. Random forests. *Mach Learn* 2001; 45: 5-32.

E10. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos T R Soc A* 2016; 374.

E11. Nixon PA, Orenstein DM, Kelsey Sf, Doershuk CF. The prognostic value of exercise testing in patients with cystic fibrosis. *N Eng J Med* 1992; 327: 1785-1788.

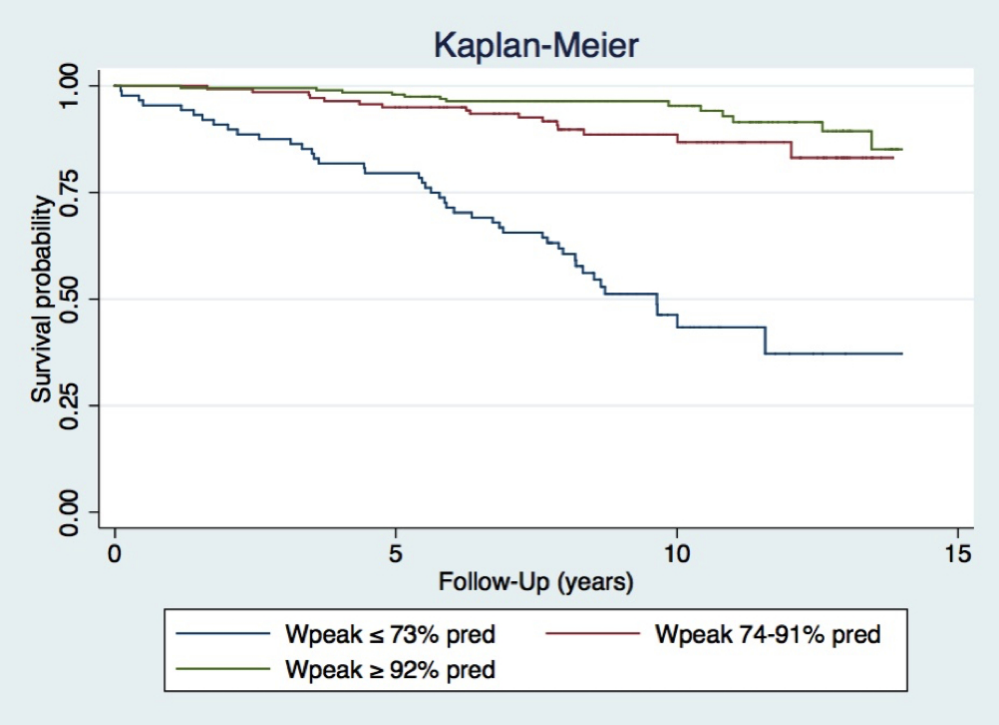

Figure E1. Kaplan-Meier survival curve for three different Wpeak groups. Wpeak, peak work rate.

AJRCCM Articles in Press. Published on 15-October-2018 as 10.1164/rccm.201806-1110OC

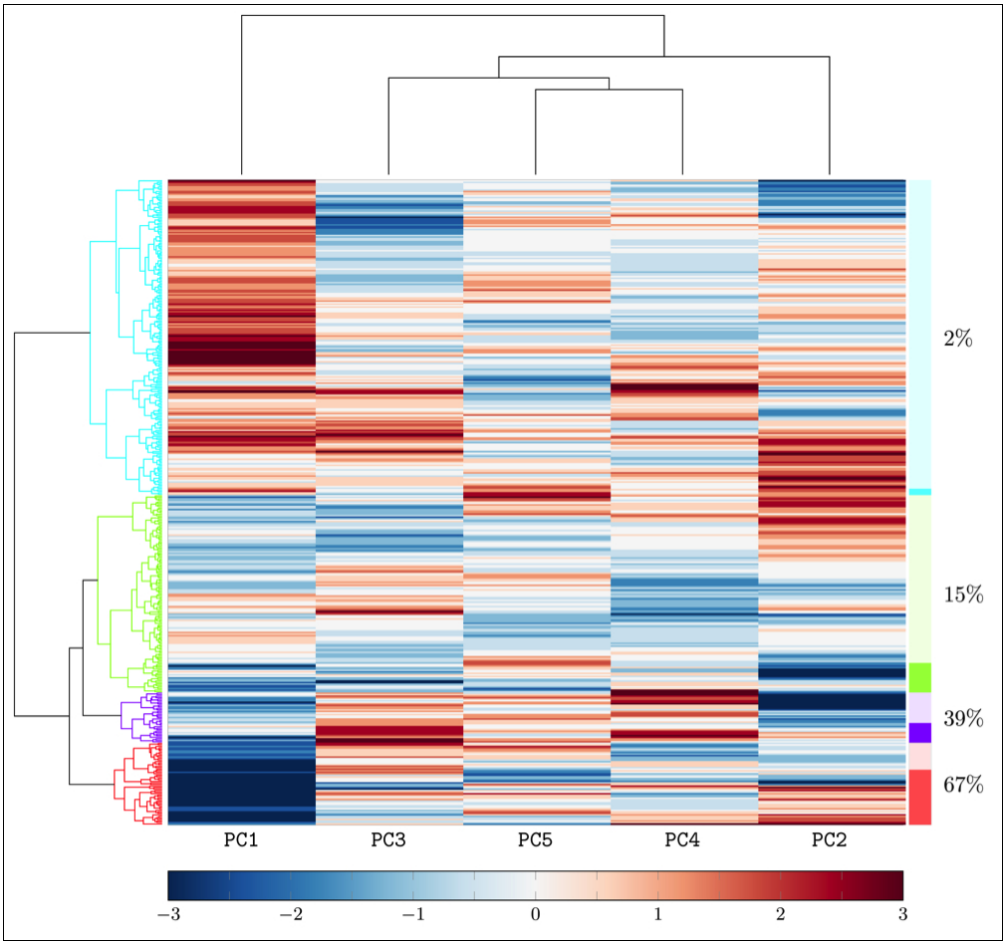

Figure E2. Heat map representing hierarchical Ward's clustering. The left color bar denotes individual subject grouping and their related cluster indicated by different colors: cluster one (Cyan, n=207), cluster two (green, n=130), cluster three (red n=54) and cluster four (purple, n=33). Column on the right: colors correspond to the clusters reported on the left bar, and for each cluster the percent of the main outcome (deaths or LTx) is reported. Horizontal bar denotes the values from each principal component (PC): dark blue represents values below the mean, white represents the mean value, and dark red represents values above the mean.