Pulmonary Outcome Prediction (POP) Tools for Cystic Fibrosis Patients

Donald R. VanDevanter, PhD[1], **Jeffrey S. Wagener, MD**[2], **David J. Pasta, MS**[3], **Eric Elkin, MPH**[3], **Joan R. Jacobs, MS**[4], **Wayne J. Morgan, MD**[5], and **Michael W. Konstan, MD**[1]

[1]Case Western Reserve University School of Medicine, Cleveland, OH

[2]University of Colorado Denver School of Medicine, Aurora, CO

[3]ICON Clinical Research, San Francisco, CA

[4]Genentech, Inc., South San Francisco, CA

[5]University of Arizona, Tucson, AZ

Summary

Rationale—Loss of lung function in patients with cystic fibrosis (CF) is associated with increased mortality and varies between individuals and over time. Predicting this decline could improve patient management.

Objectives—Develop simple pulmonary outcome prediction (POP) tools to estimate lung function at age 6 in patients aged 2–5 years (POP_{2-5}) and lung function change over a 4 year period in patients aged 6–17 years (POP_{6-17}).

Methods—Analyses were conducted using patients from the Epidemiologic Study of CF (ESCF). To be included in any analysis, patients had to have one year of clinical history recorded in ESCF prior to a clinically stable routine Index Clinic Visit (ICV). In addition to this criterion, for the POP_{2-5} tool patients had to be between 2 and 5 years old at ICV and have a second clinically stable visit with spirometric measures at age 6. For the POP_{6-17} tool, patients had to be between the ages of 6 and 17 years old at an ICV that included spirometric measures and had to have a second clinically stable visit with spirometric measures from 3 to 5 years after ICV. All patients enrolled in ESCF who met these inclusion criteria were studied. POP_{2-5} and POP_{6-17} populations were further divided into development groups (with ICV before 1/1/1998) and validation groups (with ICV after that date). Development groups were used to model forced expiratory volume in 1 s (FEV_1) percent predicted at age 6 years (for POP_{2-5}) and annualized

$FEV_1\%$ predicted change from ICV to the second visit (for POP_{6-17}) by multivariable linear regression using age, sex, weight-for-age percentile, cough, sputum production, clubbing, crackles, wheeze, sinusitis, number of exacerbations requiring intravenous antibiotics in the past year, elevated liver enzymes, pancreatic enzyme use, and respiratory tract culture status, plus height-for-age percentile (POP_{2-5}) and index FEV_1 (POP_{6-17}). Integer-based POP_{2-5} and POP_{6-17} tools created from selected variables were evaluated by Pearson correlation and then prospectively validated with separate data collected later from ESCF patients with ICV after 1/1/1998.

Main Results—POP_{2-5} and POP_{6-17} development groups included 2,709 and 6,113 patients and validation groups included 3,458 and 7,086 patients, respectively. Variables retained were weight-for-age percentile, clubbing, crackles, wheeze, number of exacerbations, and *Pseudomonas aeruginosa* culture status (both tools), daily cough (POP_{2-5}), and age, sex, and index $FEV_1\%$ predicted (POP_{6-17}). Correlation coefficients for POP_{2-5} and POP_{6-17} tools prospectively applied to validation groups were +0.32 and +0.37, respectively.

Conclusions—These simple integer-based POP algorithms employ variables available at clinic visits and can be used to predict the probability of different future pulmonary outcomes for individual patients and patient populations.

Keywords

epidemiology; predictive variables; pulmonary function outcome; decline

Introduction

Several algorithms have been described to assist clinicians in assessing the lung disease status of individuals with cystic fibrosis (CF). Algorithms such as Chrispin-Norman,[1] Brasfield,[2] Wisconsin,[3] Northern,[4] and Age-Based Severity[5] produce scores based upon chest radiographs, while Shwachman-Kulczycki[6] and NIH[7] algorithms employ radiographic, pulmonary function test (PFT), and clinical symptom data. These scoring systems primarily provide insight into lung disease stage, although some were developed to predict near-term mortality.[7] Today, no single scoring system is routinely used across a broad segment of the CF clinical community, which may be a consequence of increased CF survival. For instance, the practical value of assessing the probability of near-term mortality for all patients by NIH score[7] decreases correspondingly as the fraction of patients likely to experience near-term mortality approaches zero.

Today, there is a greater emphasis on the use of pulmonary function, and particularly forced expiratory volume in 1 s (FEV_1) and the fraction of FEV_1 compared to a reference population ($FEV_1\%$ predicted), than on radiographic scoring tools for assessing CF lung disease stage.[8] Although Rosanthal[9] suggested that reliance on FEV_1 as an indicator of health status is questionable, the measure remains an influential driver for treatment of pulmonary exacerbations[10] and is central to the regulatory approval of CF respiratory therapies.[11]

This attention to pulmonary function may explain why clinical use of respiratory therapies is associated with a patient's disease status as measured by FEV_1.[12,13] Thus the potential to

estimate an individual's probability of FEV_1 change over the next few years can provide clinical guidance for the use of interventions intended to preserve lung function in patients most at risk, and thus most likely to benefit.

Unfortunately, variability of FEV_1 over short time intervals prohibits accurate estimation of an individual's current rate of decline, and several years of data collection are required to obtain reasonable estimates of past decline rates.[14] Additionally, past decline rates are not predictive of future decline rates. For example, young patients who have lost little or no FEV_1 are at the greatest risk of decline in the subsequent 4 years of their lives.[15] In CF patients under age 6 years, traditional spirometry is difficult and may be unreliable, although it is clear that lung disease begins early in life.[16] Although a variety of procedures have been described to assess pulmonary function in children under 6, no single method is widely available or uniformly employed by a majority of CF clinicians.[17]

A recently published method using data from the Epidemiologic Study of Cystic Fibrosis (ESCF) allows clinicians to estimate the annualized change in FEV_1 over the next 4 years in individual CF patients aged 6–17 years using a multivariate equation that weights the contribution of various risk factors (e.g., weight for age, presence of physical findings, infection status).[15] The method produces an estimate of annualized future change in FEV_1% predicted by adding parameter estimates associated with categories of 11 predictive variables to an overall estimated change rate, with separate parameter estimates and overall rates provided for 3 age groups (6–8 years, 9 – 12 years, and 13 – 17 years). Although this helps to identify CF patients at high risk for lung function decline, the method may not be practical to use in the clinical setting and is not applicable to children less than 6 years old. We therefore have extended this work by developing a simple clinical scoring tool to estimate annualized change in FEV_1% predicted over the near future for CF patients aged 6–17 years that we call POP_{6-17} ("pulmonary outcome prediction"). We used a similar approach to develop a second POP tool (POP_{2-5}) to estimate FEV_1% predicted at age 6 years for CF patients aged 2–5 years.

Methods

Data were obtained from ESCF, an encounter-based, longitudinal, multicenter, prospective observational study of North American CF patients conducted from 1994 to 2005.[18] Written informed consent was obtained according to research policies at participating institutions. Baseline information including birth date and sex were recorded when the patient enrolled in the study. Pulmonary function test (PFT) results, height, weight, clinical signs and symptoms, pulmonary exacerbations requiring intravenous (IV) antibiotic therapy, and respiratory tract culture results were recorded at each patient visit. FEV_1 values were converted to percent predicted (FEV_1 % predicted) using the reference equations of Wang et al.[19] for females through age 15 and males through age 17 and Hankinson et al.[20] at older ages. Detailed explanations of how these variables were captured in the ESCF have been provided previously.[15,18]

POP scoring algorithms were constructed for patients aged 2–5 years (POP_{2-5}) and 6–17 years (POP_{6-17}). To be included in these analyses, patients had to have a year of clinical

history recorded in ESCF prior to a clinically stable routine clinic visit (Index Clinic Visit [ICV]). In addition to this criterion, for the POP_{2-5} tool, patients had to be between 2 and 5 years old at ICV and have a second clinically stable visit with spirometric measures at age 6. For the POP_{6-17} tool, patients had to be between the ages of 6 and 17 years old at an ICV that included spirometric measures and had to have a second clinically stable visit with spirometric measures from 3 to 5 years after ICV. All patients enrolled in ESCF who met these criteria were included in the study. The two populations were further divided into development (with ICV before 1/1/1998) and validation (with ICV after that date) groups. The POP_{2-5} outcome was FEV_1% predicted at age 6 ($FEV_1$6yr) obtained at a time of clinical stability during the calendar year between the patients' 6th and 7th birthdays. The POP_{6-17} outcome was the annualized rate of FEV_1% predicted change (FEV_1change) estimated by subtracting FEV_1% predicted at the ICV from FEV_1% predicted at a stable follow-up PFT (within 4 ± 1 years) and dividing by the elapsed years between measures.

POP_{2-5} and POP_{6-17} scoring algorithms were constructed from their respective development groups by iterative regression modeling with variables available at the ICV for each development group using SAS version 9.1 (SAS Institute, Inc., Cary, NC). Patient age, sex, weight-for-age percentile, daily cough, daily sputum production, clubbing, crackles, wheeze, sinusitis, number of exacerbations treated with IV antibiotics in the past year, elevated liver enzymes, pancreatic enzyme use, and respiratory tract culture status plus height-for-age percentile (for patients aged 2–5 years) and FEV_1% predicted at the ICV (for patients aged 6–17 years) were initially included as risk variables based on prior studies.[15,21] Parameter and least squares means estimates from initial regressions were used to empirically determine which variables and interactions were to be retained in subsequent iterations and whether combinations of categories or variable levels were indicated. Model reduction proceeded iteratively with sensitivity analyses used to explore alternate parameterizations. Once each model had been trimmed to a smaller number of predictive variables, integer values were assigned to each variable based on parameter estimates. This tentative score was then treated as a single variable and candidate variables were tested to see if they added significant explanatory power, in which case they were added to the score (or the integer value assigned to the variables was changed). Through this trial and error process, integer values for each variable were adjusted until resulting POP scores (obtained by summing all integers assigned to each variable present in a given patient) provided the best fit to the desired relationship between the POP score and the pulmonary outcome it was intended to predict. For the final POP_{2-5} algorithm, each score unit was intended to approximate a 2% predicted deviation of FEV_1 from 100% predicted. Therefore, more *negative* POP_{2-5} scores were intended to predict greater *negative* deviations from 100% predicted (i.e., reduced pulmonary function). For the final POP_{6-17} algorithm, each score unit was intended to represent an FEV_1change of approximately 0.5% predicted per year, with a more *negative* POP_{6-17} score intended to predict a *greater* annualized rate of decline in FEV_1% predicted over the next 4 years.

POP_{2-5} and POP_{6-17} algorithms were then prospectively validated by assessing associations between POP scores and actual $FEV_1$6yr (2- to 5-year-olds) or FEV_1change (6- to 17-year-olds) with separate patient data that had not been included in POP tool development

(validation groups) using Pearson correlation coefficients and regressions. The purpose of validating POP tools with separate patient data was to reduce possible inflation of Type I error during trial and error assignment of integers to variable categories during tool construction. This is a common approach when developing predictive models when there are many potential predictors. For the POP_{2-5} and POP_{6-17} validation groups, Pearson correlation coefficients greater than or equal to 0.06 and 0.04, respectively, are statistically significant at $P < 0.001$ due to the large sample sizes in this study. In addition, c-indices for concordance between each possible pair of unequal POP scores and observed outcomes for validation groups were determined,[22] as well as the concordance between POP scores with differences of 2 or more points and observed outcomes. The predictive power of the POP_{6-17} scoring algorithm was compared to the previously published multivariate method developed for this patient population.[15] The outcomes $FEV_1 6yr$ and $FEV_1 change$ and their associated Z-score values were recalculated using the reference equations of Stanojevic and colleagues.[23] Associations of POP scores with recalculated outcomes were assessed using Pearson correlation coefficients and regression.

Results

Development

Overall, 34,488 patients have participated in ESCF. All ESCF patients meeting *a priori* inclusion criteria (age range, enrollment period, and stable pulmonary function measures at prescribed times) were included in the current analyses. Inclusion criteria were such that a single patient could be included in multiple groups (at different ages). 1,344 patients were included in both the development and validation groups of the POP_{2-5} tool (because they had eligible ICV both before and after January 1, 1998) and 4,449 patients were included in both the development and validation groups of the POP_{6-17} tool. An additional 1,064 patients were included in both the development group for the POP_{2-5} tool and the validation group for the POP_{6-17} tool. All other patients (N=5652) were only included in one group. The POP_{2-5} development and validation groups contained 2,709 and 3,458 patients aged 2–5 years with mean ages of 3.4 (SD 1.2) and 3.3 (SD 1.2) years, respectively. The POP_{6-17} development and validation groups contained 6,113 and 7,086 patients aged 6–17 years with mean ages of 11.1 (SD 3.3) and 11.2 (SD 3.3) years, respectively. Mean FEV_1 at the ICV for POP_{6-17} development and validation groups were 84.6% predicted (SD 21.8) and 85.6% predicted (SD 21.3), respectively, with similar FEV_1 distributions (data not shown). Distributions of other predictive variables for all populations are shown in Table 1. The median number of clinic visits included in estimation of IV exacerbations in the year prior to the ICV (Table 1) was 4 for both development groups and the POP_{6-17} validation group and 5 for the POP_{2-5} validation group, with a range from 1 to 30 visits. Of those patients with a history of respiratory tract culture in the year prior to ICV, a median of 2 cultures (range 1 to 23) were used to compile microbiology variables (Table 1).

Predictive variables retained in the final POP_{2-5} and POP_{6-17} scoring algorithms included weight-for-age percentile, clubbing, crackles, wheeze, number of exacerbations treated with IV antibiotics in the past year, and *Pseudomonas aeruginosa* culture history. In addition, daily cough remained as a variable in the POP_{2-5} algorithm, and sex, age, and FEV_1%

predicted at the ICV remained as variables in the POP_{6-17} algorithm. Predictive variables and their associated integer values were incorporated into card format to allow convenience and rapid calculation of an individual's POP score during a clinic visit (Fig. 1).

Validation

$FEV_1$6yr values for the POP_{2-5} validation group averaged 96.7% predicted (median 98.2, SD 18.9), with 31.9% of $FEV_1$6yr values below 90% predicted (Fig. 2A). The relationship between POP_{2-5} scores and $FEV_1$6yr for the validation group approximated target values, with a POP_{2-5} score of 0 equal to 100.2% predicted FEV_1 (target: 100% predicted) and each POP_{2-5} score unit equal to 1.94% predicted FEV_1 (target: 2% predicted). The median POP_{2-5} score for the validation group was −1 (mean −1.8, SD 3.1), with 11.8% of patients having a POP_{2-5} score of −6 or lower (Fig. 3A). The Pearson correlation coefficient between the POP_{2-5} score and $FEV_1$6yr was 0.32 and the c-index was 0.602. FEV_1change for the POP_{6-17} validation group averaged −1.86% predicted/yr (median −1.77, SD 4.25) (Fig. 2B). More than one-third (37.0%) of patients had an annualized rate of FEV_1 decline that exceeded 3% predicted/yr, and a comparable fraction (36.7%) had annualized declines of less than 1% predicted/yr (Fig. 2B). The median POP_{6-17} score for the validation group was −4 (mean −4.09, SD 3.15) (Fig. 3B). A POP_{6-17} score of 0 represented a 0.19% predicted/yr FEV_1 change (target 0.0% predicted/yr) and each POP_{6-17} score unit equaled 0.50% predicted/yr FEV_1 change (target 0.5% predicted/yr). The Pearson correlation coefficient between the POP_{6-17} score and FEV_1change was 0.37 and the c-index was 0.631.

The associations between POP scores and FEV_1% predicted outcomes recalculated using the Stanojevic et al. method [23] were similar to those observed using the Wang et al. and Hankinson et al. reference equations (Fig. 3C and 3D). The regression of the recalculated $FEV_1$6yr as a function of the POP_{2-5} score had an intercept of 100.8% predicted FEV_1 and a slope of 1.81% predicted/unit. The Pearson correlation coefficient was 0.29. Regression of the recalculated FEV_1slope as a function of the POP_{6-17} score had an intercept of −0.23% predicted/yr and a change of 0.36% predicted/yr-unit score. The Pearson correlation coefficient was 0.28. Association between POP scores and pulmonary function endpoints recalculated as Stanojevic et al. FEV_1 Z-scores (FEV_1 Z-score at age 6 for POP_{2-5} and annualized change in FEV_1 Z-score for POP_{6-17}) had Pearson correlation coefficients of 0.29 and 0.24, respectively.

Discussion

Despite consistent incremental improvements in predicted median survival over past decades, CF remains a life-shortening disease in which loss of lung function is the primary driver of mortality.[8] For better or worse,[9] clinicians have come to rely heavily on FEV_1 as a measure of disease stage and aggressiveness,[24] need for immediate intervention,[10] and therapeutic efficacy.[11] However, a high variability in estimated future disease progression based on FEV_1 measurements creates a challenge for the CF clinician. Ideally, patients at high risk for rapid lung function decline in the near future should be candidates for more-intensive intervention, while the management of patients predicted to be more stable should be maintained as is. In this way, maintenance of the population pulmonary function can be

achieved at a lower level of health care utilization and treatment burden. Estimating a patient's actual (current) rate of lung function decline is an inherently retrospective process, leaving no opportunity for preemptive or preventative intervention.[14] When clinicians rely solely on current FEV_1 as a predictor of FEV_1 decline over the next few years, therapeutic intervention tends to be more aggressive *after* FEV_1 decline.[13] This management strategy fails to prevent initial decline and likely contributes to the observation that children and adolescents with the best lung function are at highest risk for rapid decline.[15] Previous analyses of the association of risk factors in 3-year-old children with poorer FEV_1 at age 6[21] and use of risk factors in 6- to 17-year-old children to estimate future rate of FEV_1 decline[15] have been recognized by clinicians but have not proven simple enough to find their way into routine clinical practice. We have therefore developed two simple POP (pulmonary outcome prediction) tools using an approach previously employed in other contexts[25–27] that can be used in the clinic to estimate future lung function in children and adolescents with CF (Fig. 1). The POP_{2-5} tool (for CF patients aged 2–5 years) predicts FEV_1% predicted at age 6 ($FEV_1$6yr), extending earlier analyses showing that certain risk factors (e.g., low weight-for-age percentile) at age 3 are associated with reduced $FEV_1$6yr.[21] Possible POP_{2-5} scores range from −18 to +2, with more negative scores representing worse predicted outcomes (i.e., lower FEV_1% predicted at age 6). The POP_{6-17} tool (for patients aged 6–17 years) predicts annualized change in FEV_1% predicted (FEV_1change) over the next 4 years. Possible POP_{6-17} scores range from −16 to +9, with more negative scores representing worse predicted outcomes (i.e., higher annualized rates of FEV_1% predicted decline over the next 4 years). The POP_{6-17} tool is an extension of a previous analysis that produced a linear regression equation for estimating future rate of FEV_1 decline based on risk factors.[15]

Generation of integer-based predictive tools includes a trial and error assignment of integer values to parameter estimates has the potential to introduce additional Type I error by including variables that are not actually predictive but have been included by chance. For this reason, the POP tools were modeled using patient data from a development group and then prospectively tested on a separate set of patient data (the validation group).

These POP tools are by no means the first integer-based scoring tools to be described for assessing the clinical status of CF patients. Important predecessors include radiographic scoring tools such as the Chrispin-Norman,[1] Brasfield,[2] Wisconsin,[3] Northern,[4] and Age-Based Severity[5] scores and mixed radiographic/symptom-based tools such as the Shwachman-Kulczycki[6] and NIH[7] scores. These scoring systems have been used largely as disease staging (i.e., comparative) rather than predictive tools, although some have been used to assess risks of near-term mortality.[7] No single one of these scores is in wide use today, although all have value for staging CF lung disease and following disease progression. Recently, lung disease staging using high-resolution computerized tomography (HRCT) has gained momentum,[28–30] but has yet to be widely incorporated into routine clinical practice.

Our interest in developing these tools was to provide a method for clinicians to estimate the FEV_1 of their patients in the near future, and most importantly, to identify those patients whom epidemiologic experience suggests have an increased probability of an unfavorable future FEV_1 relative to current clinical impression. An optimal CF pulmonary outcome

prediction tool should have reasonable precision while retaining enough simplicity to allow routine use in the clinic. To achieve these characteristics, a balance must be struck between the increased precision achieved through inclusion of many possible predictive variables and the reduced complexity achieved by elimination or combination of as many variables as possible. Further, *how* a variable is defined can also affect tool utility. Specifically, variables that are easily accessible in the clinic should make the tool more practical. For example, we chose "exacerbations treated with IV antibiotics in the past year" rather than defining an exacerbation using a complex algorithm composed of signs and symptoms. An unambiguous record of IV treatments should be readily accessible in a patient's chart at a clinic visit and there is no need to reconstruct signs and symptoms to determine whether episodes during the past year met clinical criteria for defining an exacerbation. Finally, because the POP tools have been developed and validated with data collected from CF caregivers at over 100 sites, inter-observer variability has been accounted for within the variance of each tool.

Clinicians should not conclude that the variables included in our POP tools are the *only* ones that might be relevant or useful for predicting FEV_1 in the future. Rather, these combined variables provide reasonable precision to predict relative probabilities of poor $FEV_1$6yr and FEV_1change outcomes. For instance, it is well-established that CFTR mutation type(s) influence CF disease course, [31,32] although outcomes are known to vary among patients with identical CFTR genotypes.[33] The variable "use of pancreatic enzymes," which can be considered a surrogate for the presence of two "severe" CFTR mutant alleles, was included as a candidate variable during modeling. However, it did not survive backward elimination during development of either POP tool, presumably because remaining variables adequately captured risks associated with different CFTR mutations. Interestingly, when we compared our POP_{6-17} tool with our more complex regression equation that included eleven potentially predictive variables, including use of pancreatic enzymes,[15] this simpler tool exhibited better precision in predicting FEV_1change for the validation group (Pearson correlation coefficient of 0.37 for the current method versus 0.26 for the previous method). Addition of CF radiographic assessment scores into POP tools might have increased tool precision. However, it would have been problematic to include a radiographic assessment as no single score is widely used by CF physicians today (for instance, less than a quarter of patients enrolled in ESCF had at least one Brasfield chest roentgenogram score[2] recorded).

For any POP score value, there is a distribution of observed patient outcomes (Fig. 3C and 3D), with some patients having outcomes worse than would be predicted by the POP score and others with outcomes better. The c-indices for the POP_{2-5} and POP_{6-17} scores were 0.602 and 0.631, respectively. The c-index is the probability of concordance between a POP score and the outcome it predicts.[22] A pair of patients with different POP scores is considered concordant if the patient with the lower POP score also experiences a worse outcome. The c-index is the fraction of all possible pairs with unequal POP scores in which the patient with the lower POP score had a worse outcome. Each pair of patients with unequal POP scores and equal outcomes is assigned a concordance of 0.5. A c-index of 0.5 would indicate a random predictive value of the POP tool; a c-index of 1.0 would be perfectly concordant. A difference of 1 POP tool point between two patients predicts a difference in outcome of only 2 points of FEV_1 % predicted (2% predicted FEV_1 at age 6

for POP$_{2–5}$ and 0.5% predicted per year over 4 years for POP$_{6–17}$), a difference that is unlikely to be considered clinically compelling. Concordance for all pairs of patients with POP score differences of at least 5 units (predicting at least a 10 point difference in FEV$_1$% predicted) was 0.692 for POP$_{2–5}$ and 0.731 for POP$_{6–17}$. Concordance increased to 0.822 and 0.828, respectively, for POP score differences of 10 or more units. Thus, differences of a few POP points between patients or groups of patients are not particularly meaningful, while larger differences are quite likely to correctly predict differences in FEV$_1$ outcomes.

Variability in outcomes likely reflects differences in the underlying biological severity of CF lung disease within the population, the extent to which supportive care has been influenced by recognition of risk factors and individual clinician practice patterns, and inter-observer variability in the assessment of clinical signs and symptoms in ESCF. Although a given POP score does not unambiguously predict a patient's outcome, and individuals with scores that differ by only one or two points may not have markedly different risks (Fig. 4), the scores can be useful to categorize individuals into groups with substantially different risks. For example, a patient with a POP$_{2–5}$ score of −7 or lower was more likely than not (52%) to have an FEV$_1$ below 90% predicted at age 6, as were >75% of patients with POP$_{2–5}$ scores of −10 or lower, whereas only one in three patients with POP$_{2–5}$ scores of −2 experienced a similar outcome, with lower probabilities for patients with higher POP$_{2–5}$ scores (Fig. 4A). Similarly, a patient with a POP$_{6–17}$ score of 0 or above had only a 17.5% risk of an annualized FEV$_1$ decline of 2.75 % predicted/year or greater, whereas those scoring between −1 and −5 had double the risk (36.1%) and those scoring −6 or below had triple the risk (52.1%) (Fig. 4B).

Although there are a number of approaches and commercial devices for the assessment of lung function in CF patients too young for traditional spirometry,[17] no single method has achieved widespread use in the clinics participating in ESCF. For this reason, a single measure of this type was not available in the ESCF and could not be incorporated into the POP$_{2–5}$ tool. As they become available, these measures could be assessed for their ability to improve pulmonary outcome prediction for young children with CF. For instance, the multiple-breath inert gas washout (MBW) technique[34] has been shown to be more sensitive than spirometry for characterizing early lung disease in children with CF.[16,17,35] MBW at age 6 could be substituted for FEV$_1$% predicted and/or incorporated as a predictive variable as one of several factors to consider in a future POP$_{2–5}$ tool.

There are well-recognized precedents for developing outcome tools to assist in the management of chronic diseases, including a tool to estimate 5-year survivorship in CF.[36,37] However, managing CF lung disease progression with predictive tools presents unique challenges. For example, the status of several predictive variables incorporated into cardiovascular health management tools (e.g., blood pressure, blood cholesterol levels, smoking status) can be improved by behavioral and/or pharmacologic intervention. Further, these interventions have been demonstrated to improve outcomes and have become cornerstones of standard disease management. In the case of CF POP tools, there may be interventions that are widely believed to affect the status of predictive variables such as crackles or cough (e.g., antibiotic therapy), but objective evidence is lacking. Similarly, recent CF therapeutic guidelines recommend some chronic therapies based on the results of

clinical trials of relatively short duration, but the impact of their long-term use has not been evaluated. [38] The small size of the CF population and extended course of CF lung disease make prospective characterization of potentially disease-modifying therapies problematic and increase the importance of retrospective analyses of large CF patient registries.[39,40] Although improvements in predicted survival over the past decade have been accompanied by an associated reduction in clinical signs and symptoms, to date there is only suggestive evidence that these changes can be attributed to more aggressive intervention.[41] Despite these caveats, the potential benefit to individual CF patients arising from an ability to predict pulmonary outcomes is clear. Recognizing a high potential for a poor outcome in the near future should motivate clinicians to intensify management.

Recently, Stanojevic et al. published an alternative method for calculating FEV_1% predicted in children and proposed use of FEV_1 Z-scores for assessing pulmonary function loss.[23] This approach has the strength of accounting for the impacts of age, height, and sex on coefficients of variation, but the method is limited to non-Hispanic white populations, and the reference equations are not yet widely used. When the POP tools reported here were assessed for their ability to predict outcomes recalculated using the Stanojevic et al. reference equations and also when expressed as FEV_1 Z-scores, they retained their predictive power.

In conclusion, we have developed two tools—POP_{2-5} and POP_{6-17}—that reduce the complexity associated with estimating the risks of poor future pulmonary outcomes in children with CF aged 2–17 years. Although clinicians may already recognize that certain predictive variables (e.g., crackles, cough, and low weight for age) are cause for concern, the POP tools *weight* the contribution of each variable for pulmonary outcome. Identifying patients within their clinics who share relatively higher risks or lower risks for poor pulmonary outcomes may help clinicians standardize patient management and could lead to more rigorous intervention for previously unrecognized higher-risk patients. Additionally, the ability to segregate patients into different groups with respect to probability of poor future lung function may prove useful for clinical researchers interested in studying higher-risk populations. Finally, these simple tools may assist clinicians in communicating the relative probability of future lung disease progression and outcomes to patients with CF and their families.

Acknowledgments

References

1. Chrispin AR, Norman AP. The systematic evaluation of the chest radiograph in cystic fibrosis. Pediatr Radiol. 1974; 2:101–5. [PubMed: 15822331]
2. Brasfield D, Hicks G, Soong S-J, Tiller RE. The chest roentgenogram in cystic fibrosis: a new scoring system. Pediatrics. 1979; 63:24–29. [PubMed: 440798]

3. Weatherly MR, Palmer CG, Peters ME, Green CG, Fryback D, Langhough R, Farrell PM. Wisconsin cystic fibrosis chest radiograph scoring system. Pediatrics. 1993; 91:488–95. [PubMed: 8424032]

4. Conway SP, Pond MN, Bowler I, Smith DL, Simmonds EJ, Joanes DN, Hambleton G, Hiller EJ, Stableforth DE, Weller P, Littlewood JM. The chest radiograph in cystic fibrosis: a new scoring system compared with the Chrispin-Norman and Brasfield scores. Thorax. 1994; 49:860– 2. [PubMed: 7940422]

5. Cleveland RH, Neish AS, Zurakowski D, Nichols DP, Wohl ME, Colin AA. Cystic fibrosis: a system for assessing and predicting progression. AJR Am J Roentgenol. 1998; 170:1067–72. [PubMed: 9530060]

6. Shwachman H, Kulczyki LL. Long term study of one hundred and five patients with cystic fibrosis. Am J Dis Child. 1958; 96:6–15.

7. Taussig LM, Kattwinkel J, Friedewald WT, di Sant'Agnese PA. A new prognostic score and clinical evaluation system for cystic fibrosis. Pediatrics. 1973; 82:380–90.

8. Cystic Fibrosis Foundation Patient Registry 2007: Annual Data Report. Bethesda, Maryland: Cystic Fibrosis Foundation; 2008.

9. Rosenthal M. How good are pulmonary function tests as an indicator of short and long term health status? Pediatr Pulmonol. 2009; S32:171–2.

10. Rabin HR, Butler SM, Wohl ME, Geller DE, Colin AA, Schidlow DV, Johnson CA, Konstan MW, Regelmann WE. Epidemiologic Study of Cystic Fibrosis. Pulmonary exacerbations in cystic fibrosis. Pediatr Pulmonol. 2004; 37:400–6. [PubMed: 15095322]

11. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP). Guideline on the clinical development of medicinal products for the treatment of cystic fibrosis. EMEA/CHMP/EWP/9147/2008. Oct. 2009http://www.ema.europa.eu/pdfs/human/comp/36313208en.pdf

12. Konstan MW, Butler SM, Schidlow DV, Morgan WJ, Julius JR, Johnson CA. Patterns of medical practice in cystic fibrosis: part II. Use of therapies. Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Pediatr Pulmonol. 1999; 28:248–54. [PubMed: 10497373]

13. Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. Chest. 2003; 123(1):20–7. [PubMed: 12527598]

14. Davis PB, Byard PJ, Konstan MW. Identifying treatments that halt progression of pulmonary disease in cystic fibrosis. Pediatr Res. 1997; 41:161–165. [PubMed: 9029632]

15. Konstan MW, Morgan WJ, Butler SM, Pasta DJ, Craib ML, Silva SJ, Stokes DC, Wohl ME, Wagener JS, Regelmann WE, Johnson CA. Scientific Advisory Group and the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. J Pediatr. 2007; 151:134–139. [PubMed: 17643762]

16. Lum S, Gustafsson P, Ljungberg H, Hülskamp G, Bush A, Carr SB, Castle R, Hoo AF, Price J, Ranganathan S, Stroobant J, Wade A, Wallis C, Wyatt H, Stocks J. London Cystic Fibrosis Collaboration. Early detection of cystic fibrosis lung disease: multiple-breath washout versus raised volume tests. Thorax. 2007; 62:341–347. [PubMed: 17121870]

17. Beydon N, Davis SD, Lombardi E, Allen JL, Arets HG, Aurora P, Bisgaard H, Davis GM, Ducharme FM, Eigen H, Gappa M, Gaultier C, Gustafsson PM, Hall GL, Hantos Z, Healy MJ, Jones MH, Klug B, Lodrup Carlsen KC, McKenzie SA, Marchal F, Mayer OH, Merkus PJ, Morris MG, Oostveen E, Pillow JJ, Seddon PC, Silverman M, Sly PD, Stocks J, Tepper RS, Vilozni D, Wilson NM. American Thoracic Society/European Respiratory Society Working Group on Infant and Young Children Pulmonary Function Testing. An official American Thoracic Society/ European Respiratory Society statement: pulmonary function testing in preschool children. Am J Respir Crit Care Med. 2007; 175:1304–1345. [PubMed: 17545458]

18. Morgan WJ, Butler SM, Johnson CA, Colin AA, FitzSimmons SC, Geller DE, Konstan MW, Light MJ, Rabin HR, Regelmann WE, Schidlow DV, Stokes DC, Wohl ME, Kaplowitz H, Wyatt MM, Stryker S. Epidemiologic Study of Cystic Fibrosis: design and implementation of a prospective, multicenter, observational study of patients with cystic fibrosis in the U.S and Canada. Pediatr Pulmonol. 1999; 28:231–241. [PubMed: 10497371]

19. Wang X, Dockery DW, Wypij D, Fay ME, Ferris BG. Pulmonary function between 6 and 18 years of age. Pediatr Pulmonol. 1993; 15:75–88. [PubMed: 8474788]

20. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med. 1999; 159:179–187. [PubMed: 9872837]

21. Konstan MW, Butler SM, Wohl ME, Stoddard M, Matousek R, Wagener JS, Johnson CA, Morgan WJ. Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Growth and nutritional indexes in early life predict pulmonary function in cystic fibrosis. J Pediatr. 2003; 142:624–630. [PubMed: 12838189]

22. Harrell, FE, Jr. Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer-Verlag; 2001. p. 247-248.

23. Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, Rosenthal M, Corey M, Lebecque P, Cole TJ. Reference ranges for spirometry across all ages: a new approach. Am J Respir Crit Care Med. 2008; 177:253–260. [PubMed: 18006882]

24. Konstan MW, Wagener JS, VanDevanter DR. Characterizing aggressiveness and predicting future progression of CF lung disease. J Cyst Fibros. 2009; 8 (Suppl 1):S15–9. [PubMed: 19460682]

25. Cooperberg MR, Pasta DJ, Elkin EP, Litwin MS, Latini DM, Du Chane J, Carroll PR. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. J Urol. 2005; 173:1938–42. [PubMed: 15879786]

26. Miller MK, Lee JH, Blanc PD, Pasta DJ, Gujrathi S, Barron H, Wenzel SE, Weiss ST. TENOR Study Group. TENOR risk score predicts healthcare in adults with severe or difficult-to-treat asthma. Eur Respir J. 2006; 28:1145–55. [PubMed: 16870656]

27. Adamson GD, Pasta DJ. Endometriosis fertility index: the new, validated endometriosis staging system. Fertil Steril. 2009 Nov 18. Epub ahead of print.

28. Robinson TE. High-resolution CT scanning: potential outcome measure. Curr Opin Pulm Med. 2004; 10:537–41. [PubMed: 15510063]

29. Brody AS, Kosorok MR, Li Z, Broderick LS, Foster JL, Laxova A, Bandla H, Farrell PM. Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. J Thorac Imaging. 2006; 21:14–21. [PubMed: 16538150]

30. Davis SD, Brody AS, Emond MJ, Brumback LC, Rosenfeld M. Endpoints for clinical trials in young children with cystic fibrosis. Proc Am Thorac Soc. 2007; 4:418–30. Review. [PubMed: 17652509]

31. Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D, Levison H, Tsui LC, Durie P. The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (delta F508). N Engl J Med. 1990; 323:1517–22. [PubMed: 2233932]

32. McKone EF, Goss CH, Aitken ML. CFTR genotype as a predictor of prognosis in cystic fibrosis. Chest. 2006 Nov; 130(5):1441–7. [PubMed: 17099022]

33. Schluchter MD, Konstan MW, Drumm ML, Yankaskas JR, Knowles MR. Classifying severity of cystic fibrosis lung disease using longitudinal pulmonary function data. Am J Respir Crit Care Med. 2006; 174:780–786. [PubMed: 16858011]

34. Kraemer R, Blum A, Schibler A, Ammann RA, Gallati S. Ventilation inhomogeneities in relation to standard lung function in patients with cystic fibrosis. Am J Respir Crit Care Med. 2005; 171:371–8. [PubMed: 15531750]

35. Gustafsson PM, Aurora P, Lindblad A. Evaluation of ventilation maldistribution as an early indicator of lung disease in children with cystic fibrosis. Eur Respir J. 2003; 22:972–979. [PubMed: 14680088]

36. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: The Reynolds Risk Score. JAMA. 2007; 297:611–619. [PubMed: 17299196]

37. Liou TG, Adler FR, Fitzsimmons SC, Cahill BC, Hibbs JR, Marshall BC. Predictive 5-year survivorship model of cystic fibrosis. Am J Epidemiol. 2001; 153:345–352. [PubMed: 11207152]

38. Flume PA, O'Sullivan BP, Robinson KA, Goss CH, Mogayzel PJ, Willey-Courand DB, Bujan J, Finder J, Lester M, Quittell L, Rosenblatt R, Vender RL, Hazle L, Sabadosa K, Marshall B. Cystic Fibrosis Foundation, Pulmonary Therapies Committee. Cystic fibrosis pulmonary guidelines:

chronic medication for maintenance of lung health. Am J Respir Crit Care Med. 2007; 176:957–969. [PubMed: 17761616]

39. Konstan MW, Schluchter MD, Xue W, Davis PB. Clinical use of ibuprofen is associated with slower FEV_1 decline in children with cystic fibrosis. Am J Respir Crit Care Med. 2007; 176:1084–1089. [PubMed: 17872492]

40. Ren CL, Pasta DJ, Rasouliyan L, Wagener JS, Konstan MW, Morgan WJ. The initiation of inhaled corticosteroid therapy in cystic fibrosis patients is associated with a slower rate of lung function decline. J Pediatr. 2008; 153:746–751. [PubMed: 18760805]

41. VanDevanter DR, Rasouliyan LH, Murphy TM, Morgan WJ, Ren CL, Konstan MW, Wagener JS. for the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Trends in the clinical characteristics of the U.S. cystic fibrosis patient population from 1995 to 2005. Pediatr Pulmonol. 2008; 43:739–744. [PubMed: 18613041]

CF POP₂₋₅
Pulmonary Outcome Prediction
Age 2 to 5 years

Pseudomonas aeruginosa Culture History

All cultures in past year negative	No cultures in past year*	Any cultures in past year positive
+1	0	-1

*- Only for patients with no culture history in the past year: culture results from the current clinic visit can be used to generate the risk score.

IV Exacerbations in Past Year

none	one	two	three +
0	-1	-3	-5

Weight for Age Percentile

50th+	<50th to 25th	<25th to 10th	<10th
+1	0	-2	-4

Daily Cough

absent	present
0	-1

Clubbing

absent	present
0	-2

Crackles

absent	present
0	-5

Wheeze

absent	present
0	-2

POP₂₋₅ Score

CF POP₆₋₁₇
Pulmonary Outcome Prediction
Age 6 to 17 years

Sex

9-12 year old male	all others
+1	0

Pseudomonas aeruginosa Culture History

All cultures in past year negative	No cultures in past year*	Any cultures in past year positive
+1	0	-1

*- Only for patients with no culture history in the past year: culture results from the current clinic visit can be used to generate the risk score.

IV Exacerbations in Past Year

none	one	two/three	four +
0	-1	-2	-5

Weight for Age Percentile

25th or higher	below 25th
0	-1

FEV_1% Predicted* Range

Age	100+	<100 to 85	<85 to 70	<70
6-8	-4	-1	+2	+6
9-12	-6	-3	-2	+1
13-17	-7	-5	-4	-1

*-Wang & Hankinson equations

Clubbing

absent	present
0	-1

Crackles

absent	present
0	-1

Wheeze

absent	present
0	+1

POP₆₋₁₇ Score

Figure 1. POP₂₋₅ and POP₆₋₁₇ score cards

Left panel, POP$_{2–5}$ score card for patients with CF aged 2–5 years. Right panel, POP$_{6–17}$ score card for patients with CF aged 6–17 years. Predictive variables are arrayed vertically within boxes on each card. Possible values for each variable (e.g., present versus absent) are arrayed horizontally and the corresponding integer values are in boxes directly below. Single integers corresponding to the variable status at the index clinic visit are recorded for each variable in the empty boxes on the right. Integers are summed to derive the POP score at the bottom.

Figure 2. Distribution of outcomes for 2- to 5-year-old and 6- to 17-year-old populations
Panel A: distribution of POP_{2-5} validation group patients by their FEV_1% predicted at age 6 years ($FEV_1$6yr). Panel B: distribution of POP_{6-17} validation group patients by their annualized rate of FEV_1% predicted change (FEV_1change).

Figure 3. Distributions of POP scores and median outcomes by POP scores
Panel A: distribution of POP_{2-5} scores for 3,458 patients in the validation group. Panel B: distribution of POP_{6-17} scores for 7,086 patients in the validation group. Panel C: median FEV_1% predicted at age 6 years ($FEV_1$6yr) values calculated using the reference equations of Wang et al.[19] and Hankinson et al.[20] (closed circles) and Stanojevic et al.[23] (open circles) for patients in Panel A by POP_{2-5} score. Bars indicate first and fourth quartile boundaries. Panel D: median annualized FEV_1% predicted change (FEV_1change) values calculated using the reference equations of Wang et al.[19] and Hankinson et al.[20] (closed circles) and Stanojevic et al.[23] (open circles) for patients in Panel B by POP_{6-17} score. Bars indicate first and fourth quartile boundaries.
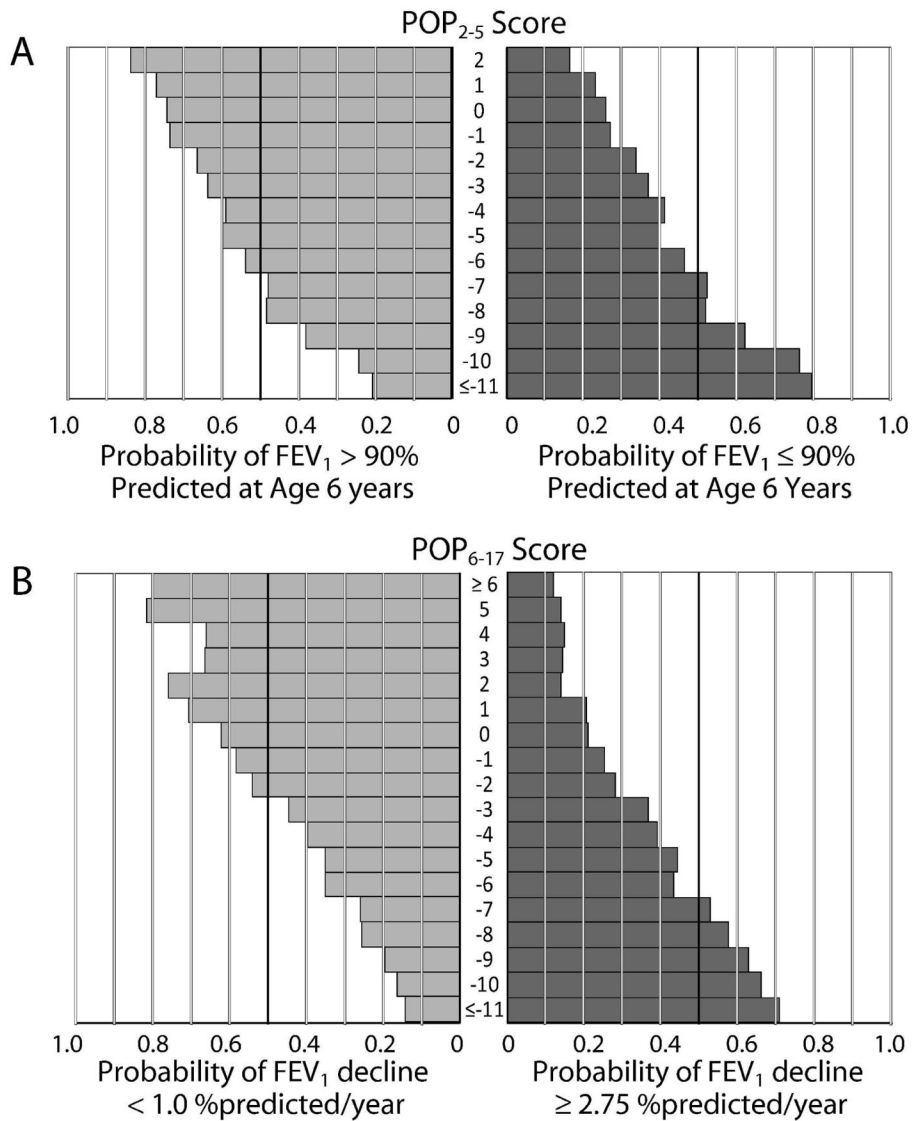
Figure 4. Probabilities of pulmonary outcomes by POP scores

Panel A: probability that a patient will experience an FEV_1 of 90% predicted (left) or <90% predicted (right) at age 6 years as a function of their POP_{2-5} score. Panel B: probability that a patient will experience an annualized decline in FEV_1 of <1% predicted/yr (left) or >2.75 % predicted/yr (right) as a function of their POP_{6-17} score.

Table 1

Patient Characteristics at the Index Clinic Visit

Variable	2–5 years old				6–17 years old			
	Development n = 2709[a,b]		Validation n = 3458[a]		Development n = 6113[c]		Validation n = 7086[b,c]	
	n	%	n	%	n	%	n	%
Female sex	1314	49	1696	49	2890	47	3343	47
Weight-for-age percentile								
<5th	305	11	357	10	1070	18	1157	16
5th to <10th	210	8	264	8	603	10	675	10
10th to <25th	556	21	661	19	1332	22	1520	21
25th to <50th	747	28	961	28	1490	24	1819	26
50th	891	33	1215	35	1618	26	1915	27
Cough								
Occasional	1157	43	1492	43	2468	40	2894	41
Daily	730	27	897	26	2673	44	2953	42
Sputum production								
Occasional	558	21	696	20	2016	33	2445	35
Daily	191	7	247	7	1514	25	1610	23
Clubbing	674	25	776	22	3394	56	3773	53
Crackles	156	6	168	5	934	15	953	13
Wheeze	97	4	92	3	297	5	282	4
Sinusitis	156	6	177	5	480	8	560	8
Number of IV exacerbations in past year								
0	2087	77	2691	78	4327	71	5129	72
1	440	16	558	16	1141	19	1232	17
2	121	4	136	4	356	6	413	6
3	34	1	46	1	151	2	167	2
4	27	1	27	1	138	2	145	2
Elevated liver function tests	40	1	72	2	106	2	166	2

Variable	2–5 years old				6–17 years old			
	Development n = 2709[a,b]		Validation n = 3458[a]		Development n = 6113[c]		Validation n = 7086[b,c]	
	n	%	n	%	n	%	n	%
Pancreatic enzymes	2601	96	3306	96	5851	96	6748	95
Microbiology in past year								
No culture	870	32	856	25	1667	27	1645	23
P. aeruginosa +	642	24	865	25	2715	44	3216	45
S. aureus +	684	25	1138	33	2021	33	2841	40
H. influenzae +	438	16	619	18	898	15	1013	14
Microbiology at index clinic visit								
No culture	1761	65	1988	57	3657	60	4127	58
P. aeruginosa +	247	9	326	9	1291	21	1419	20
S. aureus +	267	10	565	16	1054	17	1402	20
H. influenzae +	155	6	225	7	339	6	387	5

[a] - 1,344 patients included in *both* groups

[b] - 1,064 patients included in *both* groups

[c] - 4,449 patients included in *both* groups