ORIGINAL ARTICLE: CYSTIC FIBROSIS—
PEDIATRIC & ADULT

# A screening tool to identify risk for bronchiectasis progression in children with cystic fibrosis

Daan Caudri MD, PhD[1,2,3] | Lidija Turkovic PhD[1] | Nicholas H. de Klerk PhD[1] |
Tim Rosenow PhD[1] | Conor P. Murray MD[4] | Ewout W. Steyerberg PhD[5,6] |
Sarath C. Ranganathan MD, PhD[7,8,9] | Peter Sly MD, PhD[10] |
Stephen M. Stick MD, PhD[1,2] | Oded Breuer MD[1,2,11] | The AREST CF Study Group

[1]Telethon Kids Institute, The University of Western Australia, Perth, Australia

[2]Department of Respiratory Medicine, Princess Margaret Hospital, Perth, Australia

[3]Department of Pediatrics/Respiratory Medicine, Erasmus MC, Rotterdam, The Netherlands

[4]Department of Diagnostic Imaging, Princess Margaret Hospital for Children, Perth, Western Australia, Australia

[5]Department of Public Health, Erasmus MC, Rotterdam, The Netherlands

[6]Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

[7]Murdoch Children's Research Institute, Parkville, Australia

[8]Department of Respiratory Medicine, Royal Children's Hospital, Parkville, Australia

[9]Department of Paediatrics, University of Melbourne, Parkville, Australia

[10]Child Health Research Centre, The University of Queensland, Brisbane, Australia

[11]Pediatric Pulmonology and CF Unit, Department of Pediatrics, Hadassah-Hebrew University Medical Center, Jerusalem, Israel

**Correspondence**
Daan Caudri, MD, PhD, Department of Paediatric/Respiratory Medicine, Telethon Kids Institute, The University of Western

## Abstract

**Background:** The marked heterogeneity in cystic fibrosis (CF) disease complicates the selection of those most likely to benefit from existing or emergent treatments.

**Objective:** We aimed to predict the progression of bronchiectasis in preschool children with CF.

**Methods:** Using data collected up to 3 years of age, in the Australian Respiratory Early Surveillance Team for CF cohort study, clinical information, chest computed tomography (CT) scores, and biomarkers from bronchoalveolar lavage were assessed in a multivariable linear regression model as predictors for CT bronchiectasis at age 5–6.

**Results:** Follow-up at 5–6 years was available in 171 children. Bronchiectasis prevalence at 5–6 was 134/171 (78%) and median bronchiectasis score was 3 (range 0–12). The internally validated multivariate model retained eight independent predictors accounting for 37% (adjusted $R^2$) of the variance in bronchiectasis score. The strongest predictors of future bronchiectasis were: pancreatic insufficiency, repeated intravenous treatment courses, recurrent lower respiratory infections in the first 3 years of life, and lower airway inflammation. Dichotomizing the resulting prediction score at a bronchiectasis score of above the median resulted in a diagnostic odds ratio of 13 (95% confidence interval [CI], 6.3–27) with positive and negative predictive values of 80% (95% CI, 72%–86%) and 77% (95% CI, 69%–83%), respectively.

**Conclusion:** Early assessment of bronchiectasis risk in children with CF is feasible with reasonable precision at a group level, which can assist in high-risk patient selection for interventional trials. The unexplained variability in disease

Australia, 100 Roberts Rd, Subiaco, WA 6008, Australia.
Email: daan.caudri@telethonkids.org.au

progression at individual patient levels remains high, limiting the use of this model as a clinical prediction tool.

**KEYWORDS**
air trapping, bronchiectasis, lower respiratory infections, lower respiratory neutrophilic inflammation, pancreatic insufficiency, respiratory exacerbations

# 1 | INTRODUCTION

Cystic fibrosis (CF) is a severe genetic disease characterized by progressively destructive lung disease resulting in bronchiectasis and respiratory failure.[1] There is marked heterogeneity of disease expression and progression, even among patients with the same *CFTR* genotype.[2] Therefore, evaluating variables that predict accelerated disease progression, has been a focus of a considerable number of studies in patients with CF.[3]

In young preschool children, predicting accelerated disease progression is further complicated by the availability of adequate outcome measures to inform disease progression.[3,4] Spirometry, which is considered a major outcome in adults, is difficult to perform, maybe unreliable[5] and is usually normal despite significant structural lung damage as seen on computed tomography (CT).[5,6] Other measures, such as the lung clearance index (LCI), are increasingly being studied in young children, and are used as a surrogate endpoint in clinical trials.[7,8] Still, the overall mild lung disease in most preschool children with CF limits the practicality of even using LCI to routinely assess a treatment effect on lung disease. Indeed, recent trials evaluating the efficacy of novel *CFTR* modulators in preschool children present effects on nutritional status and sweat chloride measurements but not respiratory outcomes.[9–12]

Structural lung disease, and especially bronchiectasis, are a definite endpoint in lung disease progression in CF and have been shown to be associated with quality of life, mortality, and predict future lung function decline.[13–16] Thus, correctly identifying young children at increased risk of structural lung disease progression, in addition to its clinical importance, may improve patient selection for clinical trials. This will allow better assessment of treatment effects on lung health in this population.

Using data from the well-phenotyped Australian Respiratory Early Surveillance Team (AREST) CF birth cohort with follow-up until 6 years of age, we aimed to optimally predict the increase in CT bronchiectasis score to enable identification of children with the highest risk for rapid progression of structural lung disease.

# 2 | METHODS

## 2.1 | Study population

Infants diagnosed with CF after newborn screening at the CF clinics of the Princess Margaret Hospital and at the Royal Children's Hospital have been recruited in the AREST CF study. Comprehensive examinations were performed at the "annual visits" at age 3 months, 1 year, and yearly thereafter, including a bronchoscopy under general anesthesia with bronchoalveolar lavage (BAL). Since 2005 a chest CT under general anesthesia was also included in the annual follow-up in children up to the age of 5 years in Melbourne and up to the age of 6 years in Perth. The study was approved by the Ethics Committee of the Child and Adolescent Health Service (approval no. 1762EPP) and Royal Children's Hospital (approval no. 25054) and parents provided informed consent. During the study, none of the children were on CFTR modulator therapy.

## 2.2 | Data collection

At each clinical follow-up, a respiratory symptom and medication questionnaire were completed, and physical examination was recorded. BAL fluid was cultured for the detection of pathogens and analyzed for markers of inflammation (neutrophil elastase, interleukin 8, % neutrophils). Chest CT scans consisted of two volume-controlled limited slice scans obtained at end inspiration (transrespiratory pressure = 25 cmH$_2$O) and end expiration (transrespiratory pressure = 0 cmH$_2$O). Data on the number of admissions and intravenous (IV) antibiotic courses were collected from the Australian CF Data Registry.

## 2.3 | Outcome variable

CT images were scored using the CF–CT scoring method.[17] The outcome variable at end of follow-up was defined as the CF–CT bronchiectasis score at age 6 years if available, alternatively at age 5, and analyzed as a continuous score using linear regression. After the multivariable prediction model was built and validated using this continuous score, the

| | n/N (%) or median (range), N | The difference in score bronchiectasis 5–6 years | p |
|---|---|---|---|
| Predictor variables at baseline | n/N (%) | | |
| Gender (male) | 86/171 (50%) | 0.81 (−0.37 to 1.99) | 0.178 |
| Pancreatic insufficiency | 135/161 (84%) | 2.90 (1.30 to 4.50) | **<0.001** |
| Research centre (Perth) | 67/171 (39%) | 0.67 (−0.55 to 1.88) | 0.280 |
| Genetic abnormality | | | |
| Homozygous p.Phe508del | 85/169 (50%) | 0.32 (−0.88 to 1.51) | 0.602 |
| Heterozygous p.Phe508del | 75/169 (44%) | Combined as reference[a] | |
| No p.Phe508del | 9/169 (5%) | | |
| Meconium ileus (present) | 32/160 (20%) | 1.58 (0.05–3.10) | **0.042** |
| Sweat chloride (<60)[b] | 14/83 (17%) | 0.40 (0.06–0.73) | **0.020** |
| SEIFA quintile score[c] | N = 170 | −0.23 (−0.66 to 0.20) | 0.297 |
| Age-specific predictor variables | Median (IQR[d]), N | | |
| BMI age 1 | 16.9 (15.9–18.2), 111 | −0.36 (−0.77 to 0.04) | 0.078 |
| BMI age 2 | 16.3 (15.3–17.6), 107 | 0.04 (−0.39 to 0.48) | 0.838 |
| BMI age 3 | 16.3 (15.3–17.0), 107 | −0.68 (−1.19 to −0.17) | **0.010** |
| CF–CT scores | | | |
| Bronchiectasis score age 1 | 0 (0–0), 93 | 1.05 (0.50–1.59) | **<0.001** |
| Bronchiectasis score age 2 | 0 (0–1), 101 | 0.58 (0.25–0.91) | **0.001** |
| Bronchiectasis score age 3 | 1 (0–3), 120 | 0.66 (0.45–0.88) | **<0.001** |
| Air trapping score age 1 | 2 (1–5), 92 | 0.37 (0.06–0.68) | **0.019** |
| Air trapping score age 2 | 3 (1–6), 100 | 0.59 (0.37–0.80) | **<0.001** |
| Air trapping score age 3 | 2 (0–6), 120 | 0.48 (0.28–0.68) | **<0.001** |
| BAL fluid inflammation (continuous) | | | |
| % Neutrophils age 1[e] | 16.5 (7.7–34.0), 124 | 0.38 (0.03–0.73) | **0.032** |
| % Neutrophils age 2[e] | 19.5 (7.0–41.3), 134 | 0.38 (0.12–0.65) | **0.005** |
| % Neutrophils age 3[e] | 20.0 (9.3–48.0), 140 | 0.60 (0.38–0.81) | **<0.001** |
| BAL fluid inflammation (yes/no) | n/N (%) | | |
| Neutrophil elastase present age 1 | 23/124 (19%) | 1.79 (0.87–3.49) | **0.040** |
| Neutrophil elastase present age 2 | 29/137 (21%) | 1.85 (0.33–3.38) | **0.017** |
| Neutrophil elastase present age 3 | 42/141 (30%) | 2.50 (1.19–3.80) | **<0.001** |
| Interleukin 8 present age 1 | 113/124 (91%) | 2.56 (0.24–4.89) | **0.031** |
| Interleukin 8 present age 2 | 122/137 (89%) | 2.53 (0.55–4.52) | **0.013** |
| Interleukin 8 present age 3 | 129/140 (92%) | 1.71 (−0.59 to 4.00) | 0.144 |
| BAL fluid infection | | | |
| Proinflammatory pathogen age 1 | 24/124 (19%) | 1.68 (−0.00 to 3.35) | **0.050** |
| Proinflammatory pathogen age 2 | 48/137 (35%) | 1.12 (−0.19 to 2.44) | 0.094 |
| Proinflammatory pathogen age 3 | 63/144 (44%) | 2.46 (1.27–3.64) | **<0.001** |
| ≥ 2 Proinflammatory pathogens age 0–2 | 21/146 (14%) | 1.28 (0.45–2.12) | **0.003** |

**TABLE 1** General characteristics study population and univariate associations

| | n/N (%) or median (range), N | The difference in score bronchiectasis 5–6 years | p |
|---|---|---|---|
| ≥ 2 Proinflammatory pathogens age 0–3 | 45/152 (30%) | 1.68 (0.98–2.39) | **<0.001** |
| Any IV AB's last year age 1 | 29/131 (22%) | 0.94 (0.00–1.87) | **0.049** |
| Any IV AB's last year age 2 | 29/142 (20%) | 1.10 (0.17–2.02) | **0.021** |
| Any IV AB's last year age 3 | 32/147 (22%) | 1.29 (0.35–2.24) | **0.008** |

*Note*: Complete case analysis using all available data per variable. The bold values indicate $p < 0.05$.

Proinflammatory pathogens: *Pseudomonas aeruginosa, Staphylococcus aureus, Haemophilus influenzae, Streptococcus pneumoniae,* or *Aspergillus species.*

Abbreviations: AB, antibiotic; BAL, bronchoalveolar lavage; BMI, body mass index; CF, cystic fibrosis; CT, computed tomography; IQR, interquartile range; IV, intravenous; SEIFA, socioeconomic index for areas.

[a]Association analyzed as dichotomous variable homozygous p.Phe508del versus not homozygous p.Phe508del.

[b]Association analyzed as continuous variable, per 10 mmol chloride/ml increase.

[c]SEIFA defined in quintiles and association analyzed per one-step increase in quintile.

[d]IQR, 25th to 75th percentile.

[e]Association analyzed per 10% increase in percentage neutrophils.

median score at age 5–6 was used to dichotomize the predicted outcome in high versus low-risk groups.

## 2.4 | Predictor variables

Baseline predictor variables included homozygosity for *CFTR* p.Phe508del mutation, residual CFTR function, sweat chloride concentration, meconium ileus, pancreatic sufficiency, and socioeconomic status (socioeconomic index for areas 2011 census data according to postal codes at birth).[18]

Age-specific variables used as predictors included data collected at 3 months of age and at each annual follow-up thereafter until 3 years of age. These predictors included CF–CT bronchiectasis and air trapping scores, BAL results of infection and inflammation, the number of IV antibiotic courses in the year before the annual visit, and anthropometric measurements. Bacterial pathogens known to elicit a clear proinflammatory response in the lower respiratory tract (*Pseudomonas aeruginosa, Staphylococcus aureus, Haemophilus influenzae, Streptococcus pneumoniae,* and *Aspergillus species*) were also defined as a combined variable of proinflammatory pathogens.[19]

## 2.5 | Statistical analysis

Univariate associations between candidate predictor variables up to 3 years of age, and the outcome of bronchiectasis score at 5–6 years were investigated. To avoid selection bias, missing data were imputed using the "multivariate normal regression procedure" in STATA. Predictor variables were taken into account in constructing 100 multiple imputed datasets. A data-driven approach was used to select the best combination of predictors, by means of backward stepwise linear regression with likelihood

ratio statistics as a criterion for selection.[20] Akaike's information criterion was used to select predictors equivalent to $p < 0.157$.[21] The model's predictive ability was assessed using adjusted $R^2$, calibration (graphically), and Bland–Altman plot. Internal model validation was performed using a bootstrap resampling technique.[22] This enabled us to estimate the bias (i.e., optimism) caused by overfitting and to correct performance estimates as well as the regression coefficients (shrinkage) accordingly.[23] All analyses were performed using STATA 13.0.

## 3 | RESULTS

### 3.1 | Study population

A total of 171 children completed 5 or 6 years of the AREST CF program and were included in the current study. The overall proportion of missing data on the selected candidate predictor variables at baseline ($n = 8$) was 12%, and for the age-specific candidate predictor variables ($n = 10$) at Age 1 33%, at Age 2 28%, and at Age 3 25%. Additional patient characteristics and associations with bronchiectasis at 5 or 6 years of age are presented in Table 1.

### 3.2 | Outcome bronchiectasis score at 5 or 6 years

The distribution of all CF–CT subscores at Age 5–6 are given in Figure 1. One hundred and thirty-three children out of 171 (78%) had evidence of bronchiectasis (median score = 3, interquartile range = 1–8). There was a moderate to strong statistically significant correlation between bronchiectasis and all other CF–CT sub scores; air trapping score $r_s = 0.54$ ($p < 0.001$), bronchial wall thickening score $r_s = 0.63$ ($p < 0.001$), and mucus plugging score $r_s = 0.42$ ($p < 0.001$).
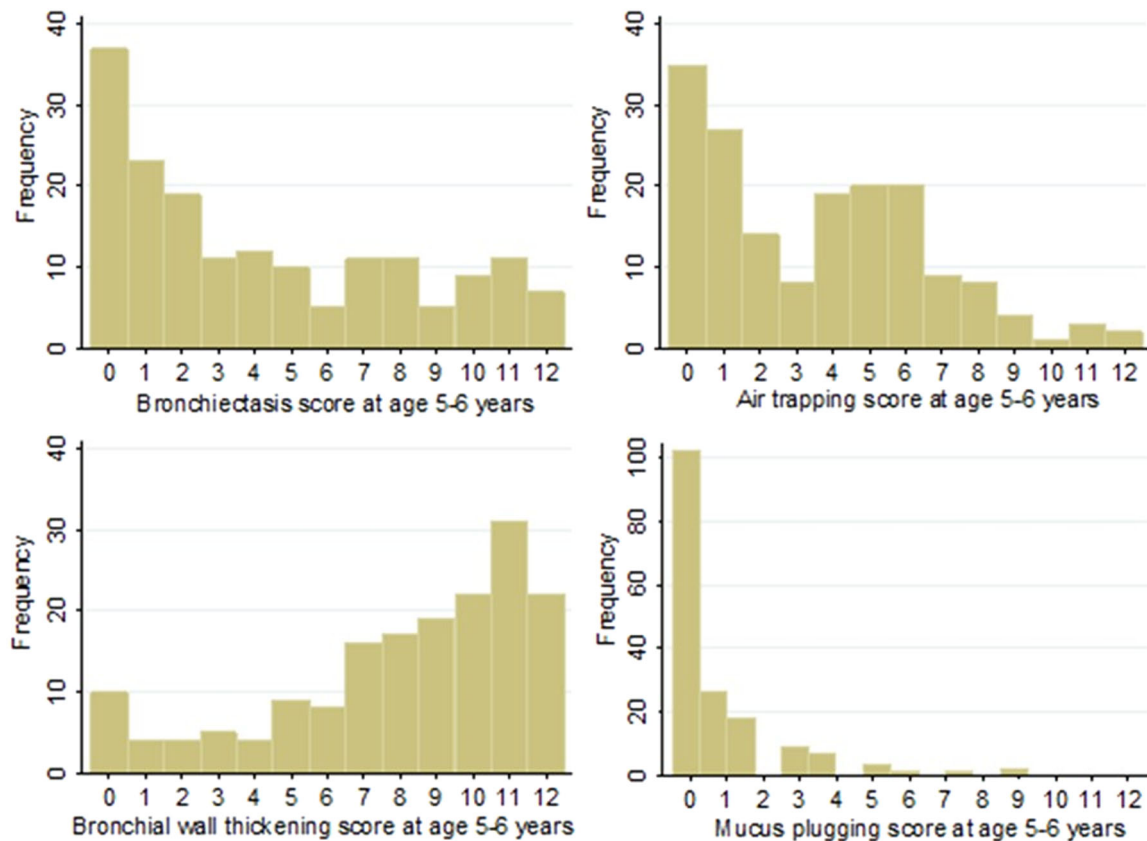
**FIGURE 1** Cystic fibrosis–computed tomography subscores distribution at the age of 5–6 (*n* = 171) [Color figure can be viewed at wileyonlinelibrary.com]

### 3.3 | Univariate analysis

Pancreatic insufficiency and meconium ileus were significant univariate predictors for later development of bronchiectasis. Homozygosity for p.Phe508del mutation was not a significant univariate predictor, using the CFTR2 database we also categorized genetic abnormalities according to their expected disease severity.[24] However, as this is mainly defined by expected pancreatic insufficiency the resulting variable did not differ substantially from our existing variable on actual pancreatic insufficiency. Sweat chloride can also be considered a functional measure of residual CFTR function and was a significant univariate predictor. As expected, CF–CT subscores were all predictive of bronchiectasis scores at 5–6. To reduce the number of variables and the risk of collinearity, only bronchiectasis and air trapping (which showed the strongest association) were included in the imputed multivariate model. Other variables found to be significant predictors of bronchiectasis scores at 5–6 are presented in Table 1.

### 3.4 | Multivariable analysis and model construction

Eight predictors were retained in the model: reduced body mass index (BMI), pancreatic insufficiency, homozygosity for p.Phe508del, two or more IV courses in the first 3 years of life, two or more lower respiratory

proinflammatory infections in the first 3 years of life, CF–CT bronchiectasis, and air trapping scores at 3 years of age and percentage of neutrophils in BAL fluid at 3 years of age (Table 2). Neutrophil elastase did not remain an independent predictor in the multivariable model despite its strong univariate association, reflecting positive correlations with other predictors (i.e., the percentage neutrophils in BAL [*p* < 0.001]).

### 3.5 | Model validation and assessment

The discriminative power of the imputed multivariate model represented by adjusted $R^2$ was 40%. Figure 2 graphically shows the predicted and actual CF–CT bronchiectasis score at 5–6 for children with complete and incomplete data. Bootstrap validation techniques to correct for overoptimism caused by overfitting estimated a validated adjusted $R^2$ to be 37%. Coefficients for individual predictors are also reported after shrinkage by bootstrap internal validation to prevent overoptimistic predictions in new populations (Table 2).

To further assess the model's predictive ability in individual patients, a Bland–Altman plot was constructed (Figure 3), showing good agreement on average between the predicted and the observed bronchiectasis scores. However, the 95% limits of agreement ranging from minus to plus 5 are wide. Furthermore, the plot seems skewed with a relative overestimation of bronchiectasis scores in children

**TABLE 2** Multivariate prediction model for CF–CT bronchiectasis score at 5–6

| Predictor variable | The difference in bronchiectasis score (CI95%) before internal validation[a] | The difference in bronchiectasis score (CI95%) after internal validation[a] | p |
|---|---|---|---|
| 1 CF–CT subscores age 3 | | | |
| 1a Bronchiectasis (Bx) score | 0.35 (0.12–0.58) | 0.25 (0.09–0.41) | 0.003 |
| 1b Air trapping (AT) score | 0.18 (−0.04 to 0.39) | 0.13 (−0.03 to 0.28) | 0.107 |
| 2 BAL fluid % neutrophils[b] age 3 | 0.33 (0.10–0.56) | 0.24 (0.08–0.41) | 0.005 |
| 3 ≥2 Proinflammatory pathogens age 0–3 | 1.19 (−0.03 to 2.42) | 0.86 (−0.02 to 1.74) | 0.056 |
| 4 ≥2 IV AB courses age 0–3 | 1.27 (−0.06 to 2.59) | 0.91 (−0.04 to 1.86) | 0.060 |
| 5 Pancreatic insufficiency | 1.55 (−0.04 to 3.13) | 1.11 (−0.03 to 2.26) | 0.056 |
| 6 BMI age 3 | −0.28 (−0.62 to 0.06) | −0.20 (−0.45 to 0.04) | 0.105 |
| 7 p.Phe508del homozygosity | −0.86 (−1.98 to 0.27) | −0.62 (−1.43 to 0.19) | 0.133 |
| Adjusted $R^2$ | 40% | 37% | |

Validated prediction model for Bx score at age 5 or 6 years:

$4.60 + 0.25 \times \text{Bx-score}_{3 \text{ years}} + 0.13 \times \text{AT-score}_{3 \text{ years}} + 0.24 \times \text{BAL \%-neutrophils}_{3 \text{ years}} + 0.86 \times \text{N proinflammatory pathogens}_{0-3 \text{ years}} + 0.91 \times \text{N IV AB courses}_{0-3 \text{ years}} + 1.11 \times \text{pancreatic insufficiency} - 0.20 \times \text{BMI}_{3 \text{ years}} - 0.62 \times \text{Delta508 homozygosity}$

*Note*: Proinflammatory pathogens: *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Haemophilus influenzae*, *Streptococcus pneumoniae*, or *Aspergillus species*.

Abbreviations: AB, antibiotic; AT, airtrapping; BAL, bronchoalveolar lavage; BMI, body mass index; BX, bronchiectasis; CI, confidence interval; CF, cystic fibrosis; CT, computed tomography; IV, intravenous.

[a]Internal validation was performed using the bootstrapping method.

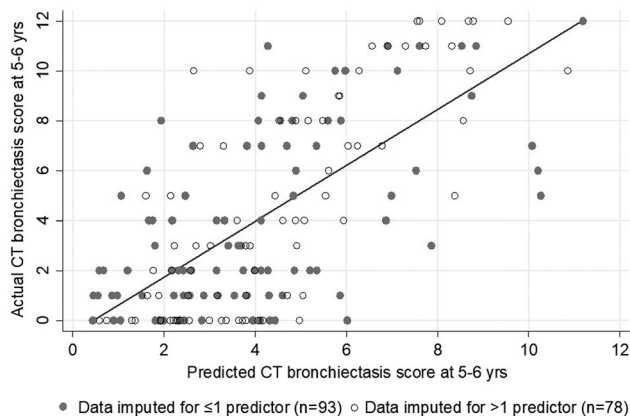[b]Association analyzed per 10% increase in percentage neutrophils.



**FIGURE 2** Predicted versus actual computed tomography (CT) bronchiectasis scores at 5–6 years (*n* = 171). All 171 children have complete data on actual CT score at Age 5–6 years. Closed circles represent children with complete data on all predictor variables (*n* = 47) or only one predictor variable missing (*n* = 46). Open circles represent children with ≥1 predictor variable missing (*n* = 78). For the figure, missing predictors were imputed with the averaged value of the 100 imputations
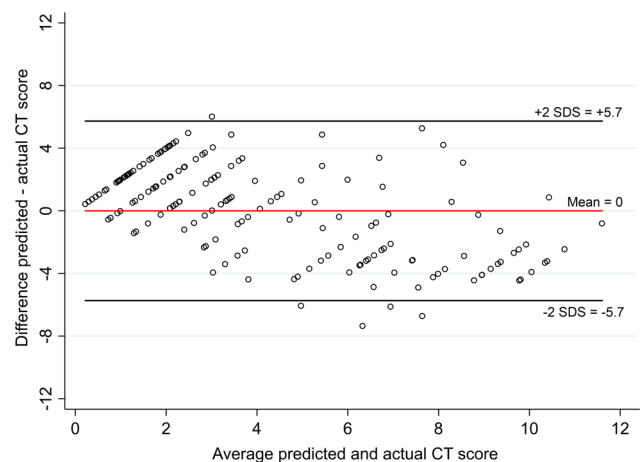


**FIGURE 3** A Bland–Altman plot showing the agreement between the predicted and actual computed tomography (CT) bronchiectasis scores at 5–6 years (*n* = 171) [Color figure can be viewed at wileyonlinelibrary.com]

**TABLE 3** Impact of predictors at Age 3 on bronchiectasis score at 5–6 in a multivariable validated model

| Predictors (continuous) | Mean | 25th percentile | Median | 75th percentile | The difference in predicted BE score at 5–6 for 25th versus 75th percentile |
|---|---|---|---|---|---|
| Bronchiectasis score at age 3 | 2.2 | 0 | 1.1 | 3.1 | +0.79 |
| Air trapping score at age 3 | 3.2 | 1 | 2.7 | 4.9 | +0.50 |
| BAL fluid % neutrophils age 3 | 31.5 | 10.3 | 26.0 | 49.0 | +0.93 |
| BMI age 3 | 16.1 | 15.2 | 16.2 | 17.0 | −0.37 |
| Predictors (dichotomous) | | % Yes | % No | | The difference in predicted Bx score at 5–6 for yes versus no |
| p.Phe508del homozygosity | | 50.0 | 50.0 | | +0.50 |
| Pancreatic insufficiency | | 84.9 | 15.1 | | +1.11 |
| ≥2 Proinflammatory pathogens age 0–3 | | 38.5 | 61.5 | | +0.86 |
| ≥2 IV AB courses age 0–3 | | 26.5 | 73.5 | | +0.91 |
| Outcome | Mean | 25th Percentile | Median | 75th Percentile | |
| Bx score at 5–6 | 4.3 | 1.0 | 3.0 | 8.0 | |

*Note*: Proinflammatory pathogens: *Pseudomonas aeruginosa, Staphylococcus aureus, Haemophilus influenzae, Streptococcus pneumoniae*, or *Aspergillus species*.

Abbreviations: AB, antibiotic; BAL, bronchoalveolar lavage; BX, bronchiectasis; BMI, body mass index; IV, intravenous.

with lower actual scores and underestimation of scores in children with higher actual scores. This would limit the clinical utility of the rule in individual patients and is an explanation for the relatively wide limits of agreement.

## 3.6 | The impact of different variables from the prediction tool

To appreciate the clinical impact of the adjusted and validated predictors, their effect size on the outcome of bronchiectasis at 5 or 6 years of age was estimated comparing the 25th versus the 75th percentile for each variable for continuous variables and presence or absence for dichotomic variables (Table 3). This shows that pancreatic insufficiency, two or more IV treatment courses and two or more lower respiratory infections in the first 3 years of life and percentage of neutrophils in BAL fluid at age 3 years, are the strongest predictors of bronchiectasis at 5 or 6 years of age. Each with an impact that is larger than that of pre-existing bronchiectasis at age 3.

## 3.7 | Use of the internally validated model

Dichotomizing the CT scores assist in identifying groups of patients with a higher risk of developing more severe bronchiectasis. This could be of particular interest when selecting patients for an interventional trial in which prevention of bronchiectasis is the primary objective. Thus, we assessed the ability of the model to identify subjects with a CF–CT bronchiectasis score above the median (CF–CT bronchiectasis score > 3) at 5–6 years of age.

Of the 85 children (50% of the total group) with a predicted bronchiectasis score of above the median, 68 children had an actual bronchiectasis score of >3 at 5–6 years. This resulted in positive and negative predictive values of 80% (95% confidence interval [CI], 72%-86%) and 77% (95% CI, 69%-83%), respectively, with a diagnostic odds ratio of 13 (95% CI, 6.3–27).

## 4 | DISCUSSION

Using data from the prospective longitudinal AREST CF cohort 8 clinical variables present at 3 years of age were selected to estimate the risk of progression of bronchiectasis at 5–6 years. The presence of pancreatic insufficiency, lower respiratory inflammation, and recurring need for treatment with IV antibiotics (recurrent infections and significant exacerbations) were the strongest predictors of bronchiectasis development in the subsequent 3 years. These variables could be used for the early identification of a group of children at high risk for lung disease progression. Disease prediction in individual patients to guide day-to-day clinical decisions remains difficult, despite the use of a comprehensive model.

Factors associated with worse CF lung disease have been previously reported. Most of these studies predicted CF sequelae in adolescence or adulthood[13,25–29] and only a minority have included young children.[4,30] From these studies, several risk factors for disease progression have been uniformly recognized. Forced expiratory volume in the 1 s at baseline is the most consistently reported predictor of future lung disease[4,13,25,26,28–30] but is unavailable for use in young children. Other classic predictors of future lung disease, more relevant for young children include *Pseudomonas aeruginosa*

colonization,[4,25,27,30] low BMI,[4,25,27–30] pancreatic insufficiency,[25,27] pulmonary exacerbations,[4,28,30] and structural lung disease.[14,16,31]

In our analysis, due to the low prevalence of *pseudomonas* in the first 3 years of life[32] *pseudomonas* infections were considered in the total number of proinflammatory pathogens from BAL and not included as a separate variable. Our data show that in this very young age group, the presence of proinflammatory pathogens (not uniquely *pseudomonas*), contribute to the development of lung damage, as has also been recently reported.[33] This suggests that early proinflammatory infections should be a target for early intervention.

Our data also support the relevance of recurrent pulmonary exacerbations as a predictor of worse lung disease and suggest that even in infancy the need for IV treatment is associated with worse structural lung damage at least up to 5–6 years of age.

As expected, we found CT bronchiectasis score at 3 years to be highly predictive of bronchiectasis at the age of 5–6 years. We also show that air trapping CT score at 3 years is predictive of later bronchiectasis, independent of concomitant CT bronchiectasis score, supporting the hypothesis that air trapping is an early consequence of inflammation and obstruction of the smaller airways, eventually leading to irreversible structural lung damage.[34] The association between structural lung damage at 3 years and structural disease at 5 or 6 years is obvious. However, our data show that markers of both inflammation and infection outweigh the effect size of early structural lung disease on predicted bronchiectasis at 5-6 years. On the basis of our model, a child with pancreatic insufficiency, high BAL neutrophils, and recurrent infections (positive BAL cultures and recurrent IV courses) is expected to have twice as much progression until the age of 5–6, regardless of pre-existing structural abnormalities on CT scan. This implies that an effective intervention in a selected group of high-risk patients could be used to improve lung disease outcomes regardless of the presence or absence of early structural abnormalities at baseline.

In the current study, using the aforementioned clinical parameters, disease biomarkers, and imaging scores, we have successfully developed a model to identify those at risk for a more rapid progression of their lung disease. One previous study by VanDevanter et al.,[4] in preschool children also developed a tool to predict school-age outcome (lung function at 6 years of age) using clinical variables. Despite this different outcome, the selected predictor variables were rather similar and differences were likely explained by the different predictor variables that were available in the two studies. For example, a physical exam (clubbing) was included only in the study by VanDevanter et al.,[4] but CT scores and BAL outcomes were not. Importantly, the score in that study showed only a weak correlation in the validation group.[4] Similarly, after validation, our detailed model including recurrent invasive assessments in all children was still expected to explain only 37% of the variance in bronchiectasis progression. Both the current and VanDevanter et al.[4] studies show us that the developed scores may allow identification of high-risk groups of individuals but, at the same time, that despite the use of comprehensive and prospectively collected data, a large proportion of variability in disease progression in individual patients cannot be accurately predicted at baseline. Other factors not evaluated in the current study, such as treatment adherence and differences in management can possibly explain at least part of this variability. Thus, clinicians and researchers should take note of the most important and consistent predictors for future disease progression and can use them for group-based follow-up schedules, or for the selection of high-risk subgroups in intervention trials. It should at the same time be realized that significant lung disease may still develop in some children with favorable prognostic indicators possibly leading to significant bronchiectasis (CF–CT scores of 8 and above, as seen in a quarter of children from our cohort).

As previously mentioned, given the overall mild lung disease in preschool children with CF recent clinical trials evaluating *CFTR* modulators did not assess respiratory indices as primary outcomes.[9–12] As bronchiectasis in CF patients is in general non-reversible and in school-aged children has been shown to consistently predict relevant clinical outcomes.[14,16,31,35] The use of a model, such as the one presented in this study may aid in preschool patient selection for clinical trials and allow assessment of a relevant respiratory outcome.

The strengths of our study lie in the prospective longitudinal design of the AREST CF cohort, the use of CT as a sensitive indicator of early lung damage, and the 99% participation in the AREST CF annual surveillance program. However, several limitations should be mentioned: first and foremost, differences in CF patient populations or management between CF centers in Australia and other parts of the world might limit generalizability, although registry data suggest that the AREST CF cohort are representative of newborn screened populations elsewhere. Second, the use of BAL biomarkers of neutrophilic inflammation and chest CT scores, are not available in many centers, limiting the full use of our prediction score. However, we also showed that simple clinical risk factors, not requiring chest CTs or BALs (pancreatic insufficiency, BMI, pulmonary exacerbations, and infections) may identify patients at high risk for more severe disease and help target treatment. Third, it should be noted that even though we performed internal validation of the model, ideally the model should be externally validated before use in clinical care.

In conclusion, our data suggest that in young children with CF, identifying individuals at high risk for disease progression is possible from 3 years of age. The use of risk assessment can be an important tool to help select high-risk children for interventional trials. However, our model had only modest ability to predict structural lung disease progression in individual patients, underlying the need to find novel and more accurate biomarkers to better monitor CF lung disease in clinical care.

## AUTHOR CONTRIBUTIONS

Please set as per style here.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

*Peter Sly* http://orcid.org/0000-0001-6305-2201

*Oded Breuer* http://orcid.org/0000-0002-4775-9095

## REFERENCES

1. Cantin A. Cystic fibrosis lung inflammation: early, sustained, and severe. *Am J Respir Crit Care Med*. 1995;151(4):939-941.
2. Drumm ML, Ziady AG, Davis PB. Genetic variation and clinical heterogeneity in cystic fibrosis. *Annu Rev Pathol*. 2012;7:267-282.
3. Breuer O, Caudri D, Stick S, Turkovic L. Predicting disease progression in cystic fibrosis. *Expert Rev Respir Med*. 2018;12(11):905-917.
4. VanDevanter DR, Wagener JS, Pasta DJ, et al. Pulmonary outcome prediction (POP) tools for cystic fibrosis patients. *Pediatr Pulmonol*. 2010;45(12):1156-1166.
5. Lum S, Gustafsson P, Ljungberg H, et al. Early detection of cystic fibrosis lung disease: multiple-breath washout versus raised volume tests. *Thorax*. 2007;62(4):341-347.
6. Tiddens HA. Detecting early structural lung damage in cystic fibrosis. *Pediatr Pulmonol*. 2002;34(3):228-231.
7. Ratjen F, Klingel M, Black P, et al. Changes in lung clearance index in preschool-aged patients with cystic fibrosis treated with ivacaftor (GOAL): a clinical trial. *Am J Respir Crit Care Med*. 2018;198(4):526-528.
8. Subbarao P, Stanojevic S, Brown M, et al. Lung clearance index as an outcome measure for clinical trials in young children with cystic fibrosis. A pilot study using inhaled hypertonic saline. *Am J Respir Crit Care Med*. 2013;188(4):456-460.
9. Davies JC, Cunningham S, Harris WT, et al. Safety, pharmacokinetics, and pharmacodynamics of ivacaftor in patients aged 2-5 years with cystic fibrosis and a CFTR gating mutation (KIWI): an open-label, single-arm study. *Lancet Respir Med*. 2016;4(2):107-115.
10. McNamara JJ, McColley SA, Marigowda G, et al. Safety, pharmacokinetics, and pharmacodynamics of lumacaftor and ivacaftor combination therapy in children aged 2-5 years with cystic fibrosis homozygous for F508del-CFTR: an open-label phase 3 study. *Lancet Respir Med*. 2019;7(4):325-335.
11. Rosenfeld M, Cunningham S, Harris WT, et al. An open-label extension study of ivacaftor in children with CF and a CFTR gating mutation initiating treatment at age 2-5years (KLIMB). *J Cyst Fibros*. 2019;18(6):838-843.
12. Rosenfeld M, Wainwright CE, Higgins M, et al. Ivacaftor treatment of cystic fibrosis in children aged 12 to <24 months and with a CFTR gating mutation (ARRIVAL): a phase 3 single-arm study. *Lancet Respir Med*. 2018;6(7):545-553.
13. Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. *N Engl J Med*. 1992;326(18):1187-1191.
14. Sanders DB, Li Z, Brody AS, Farrell PM. Chest computed tomography scores of severity are associated with future lung disease progression in children with cystic fibrosis. *Am J Respir Crit Care Med*. 2011;184(7):816-821.
15. Schluchter MD, Konstan MW, Davis PB. Jointly modelling the relationship between survival and pulmonary function in cystic fibrosis patients. *Stat Med*. 2002;21(9):1271-1287.
16. Turkovic L, Caudri D, Rosenow T, et al. Structural determinants of long term functional outcomes in young children with cystic fibrosis. *Eur Respir J*. 2020;55.
17. Brody AS, Kosorok MR, Li Z, et al. Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. *J Thorac Imag*. 2006;21(1):14-21.
18. Socio-economic indexes for areas (SEIFA). The Statistics Section of the Australian Bureau of Statistics. 2020. Accessed April 1, 2018.
19. Gangell C, Gard S, Douglas T, et al. Inflammatory responses to individual microorganisms in the lungs of children with cystic fibrosis. *Clin Infect Dis*. 2011;53(5):425-432.
20. Steyerberg EW, Eijkemans MJ, Harrell FE.Jr., Habbema JD Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making*. 2001;21(1):45-56.
21. Steyerberg EW, Eijkemans MJ, Harrell FE.Jr., Habbema JD Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med*. 2000;19(8):1059-1079.
22. Heymans MW, Buuren S, Knol DL, van Mechelen W, de van Vet HC. Variable selection under multiple imputation using the bootstrap in a prognostic study. *BMC Med Res Methodol*. 2007;7:33.
23. Steyerberg EW. Book chapter: overfitting and optimism in prediction models. *Clinical Prediction Models—A Practical Approach to Development, Validation, and Updating*. Springer-Verlag; 2009:83-99.
24. The clinical and functional translation of CFTR (CFTR2). US CF Foundation, The Hospital for Sick Children, Johns Hopkins University. http://cftr2.org. Accessed November 2016.
25. Aaron SD, Stephenson AL, Cameron DW, Whitmore GA. A statistical model to predict one-year risk of death in patients with cystic fibrosis. *J Clin Epidemiol*. 2015;68(11):1336-1345.
26. Hayllar KM, Williams SG, Wise AE, et al. A prognostic model for the prediction of survival in cystic fibrosis. *Thorax*. 1997;52(4):313-317.
27. Kerem E, Viviani L, Zolin A, et al. Factors associated with FEV1 decline in cystic fibrosis: analysis of the ECFS patient registry. *Eur Respir J*. 2014;43(1):125-133.
28. McCarthy C, Dimitrov BD, Meurling IJ, Gunaratnam C, McElvaney NG. The CF-ABLE score: a novel clinical prediction rule for prognosis in patients with cystic fibrosis. *Chest*. 2013;143(5):1358-1364.
29. Hulzebos EH, Bomhof-Roordink H, de van Weert-van Leeuwen PB, et al. Prediction of mortality in adolescents with cystic fibrosis. *Med Sci Sports Exerc*. 2014;46(11):2047-2052.
30. Emerson J, Rosenfeld M, McNamara S, Ramsey B, Gibson RL. *Pseudomonas aeruginosa* and other predictors of mortality and morbidity in young children with cystic fibrosis. *Pediatr Pulmonol*. 2002;34(2):91-100.

31. Loeve M, Hop WC, de Bruijne M, et al. Chest computed tomography scores are predictive of survival in patients with cystic fibrosis awaiting lung transplantation. *Am J Respir Crit Care Med*. 2012; 185(10):1096-1103.

32. Breuer O, Schultz A, Turkovic L, et al. Changing prevalence of lower airway infections in young children with cystic fibrosis. *Am J Respir Crit Care Med*. 2019;200(5):590-599.

33. Taylor SL, Leong LEX, Ivey KL, et al. Total bacterial load, inflammation, and structural lung disease in paediatric cystic fibrosis. *J Cyst Fibros*. 2020;19:923-930.

34. Tiddens HA, Donaldson SH, Rosenfeld M, Pare PD. Cystic fibrosis lung disease starts in the small airways: can we treat it more effectively? *Pediatr Pulmonol*. 2010;45(2):107-117.

35. Brody AS, Sucharew H, Campbell JD, et al. Computed tomography correlates with pulmonary exacerbations in children with cystic fibrosis. *Am J Respir Crit Care Med*. 2005;172(9): 1128-1132.