Using chest computed tomography and unsupervised machine learning for predicting and evaluating response to lumacaftor–ivacaftor in people with cystic fibrosis

Alienor Campredon[1,2], Enzo Battistella[3,4], Clémence Martin[2,5,6], Isabelle Durieu[6,7,8], Laurent Mely[9], Christophe Marguet[10], Chantal Belleguic[11], Marlène Murris-Espin[12], Raphaël Chiron[13], Annlyse Fanton[14], Stéphanie Bui[15], Martine Reynaud-Gaubert[16], Philippe Reix[17,18], Trieu-Nghi Hoang-Thi[1], Maria Vakalopoulou[3,4], Marie-Pierre Revel[1,2], Jennifer Da Silva[5,19], Pierre-Régis Burgel[2,5,6] and Guillaume Chassagnon[1,2] on behalf of the French Cystic Fibrosis Reference Network Study Group

[1]Radiology Dept, Hôpital Cochin, AP-HP Centre Université de Paris, Paris, France. [2]Université de Paris, Paris, France. [3]OPIS - OPtimisation Imagerie et Santé, Inria Saclay, Palaiseau, France. [4]MICS - Mathématiques et Informatique pour la Complexité et les Systèmes, CentraleSupelec, Gif-sur-Yvette, France. [5]Respiratory Medicine and Cystic Fibrosis National Reference Center, Cochin Hospital, AP-HP, Paris, France. [6]ERN-Lung CF Network, Frankfurt am Main, Germany. [7]Centre de Référence Adulte de la Mucoviscidose, Service de Médecine Interne, Hospices Civils de Lyon, Pierre Bénite, France. [8]Research on Healthcare Performance RESHAPE, Inserm U1290, Université Claude Bernard Lyon 1, Lyon, France. [9]Hôpital Renée Sabran, Cystic Fibrosis Center, Giens, France. [10]Pediatric Respiratory Disease and Cystic Fibrosis Center, Hospital UNIROUEN, Inserm EA 2656, Rouen University Hospital, Normandie Université, Rouen, France. [11]Centre de Ressources et de Compétences de la Mucoviscidose Adulte, Centre Hospitalier Universitaire de Rennes, Rennes, France. [12]Cystic Fibrosis Center, Service de Pneumologie, Pôle des Voies Respiratoires, Hôpital Larrey, CHU de Toulouse, Toulouse, France. [13]Cystic Fibrosis Center, Hôpital Arnaud de Villeneuve, Centre Hospitalier Universitaire de Montpellier, Montpellier, France. [14]Dept of Pulmonary Medicine, Cystic Fibrosis Resource and Competence Centre for Adults, Dijon University Hospital, Dijon, France. [15]Pediatric Respiratory Disease and Cystic Fibrosis Center and CIC 1401, CHU de Bordeaux, Bordeaux, France. [16]Dept of Respiratory Medicine and Lung Transplantation, Aix Marseille Universitaire, APHM, Hôpital Nord, Marseille, France. [17]UMR 5558 CNRS, Equipe EMET, Université Claude Bernard Lyon 1, Lyon, France. [18]Cystic Fibrosis Center, Hospices Civils de Lyon, Lyon, France. [19]URC-CIC Paris Descartes Necker Cochin, AP-HP, Hôpital Cochin, Paris, France.

Corresponding author: Pierre-Régis Burgel (pierre-regis.burgel@aphp.fr)

Shareable abstract (@ERSpublications)
1-year treatment with lumacaftor–ivacaftor is associated with visual improvement of bronchial abnormalities on chest CT. CT scans can help to identify patients with a higher probability of lung function improvement under lumacaftor–ivacaftor. https://bit.ly/3FUrUXv

Cite this article as: Campredon A, Battistella E, Martin C, *et al*. Using chest computed tomography and unsupervised machine learning for predicting and evaluating response to lumacaftor–ivacaftor in people with cystic fibrosis. *Eur Respir J* 2022; 59: 2101344 [DOI: 10.1183/13993003.01344-2021].

Received: 11 May 2021
Accepted: 12 Oct 2021

Abstract

Objectives Lumacaftor–ivacaftor is a cystic fibrosis transmembrane conductance regulator (CFTR) modulator known to improve clinical status in people with cystic fibrosis (CF). The aim of this study was to assess lung structural changes after 1 year of lumacaftor–ivacaftor treatment and to use unsupervised machine learning to identify morphological phenotypes of lung disease that are associated with response to lumacaftor–ivacaftor.

Methods Adolescents and adults with CF from a French multicentre real-world prospective observational study evaluating the first year of treatment with lumacaftor–ivacaftor were included if they had pre-therapeutic and follow-up chest computed tomography (CT) scans available. CT scans were visually scored using a modified Bhalla score. A k-means clustering method was performed based on 120 radiomics features extracted from unenhanced pre-therapeutic chest CT scans.

Results In total, 283 patients were included. The Bhalla score significantly decreased after 1 year of lumacaftor–ivacaftor (−1.40±1.53 points compared with pre-therapeutic CT, p<0.001). This finding was related to a significant decrease in mucus plugging (−0.58±0.88 points, p<0.001), bronchial wall thickening (−0.35± 0.62 points, p<0.001) and parenchymal consolidations (−0.24±0.52 points, p<0.001). Cluster analysis identified three morphological clusters. Patients from cluster C were more likely to experience an increase in per cent predicted forced expiratory volume in 1 s (FEV_1 % pred) ⩾5% under lumacaftor–ivacaftor than those in the other clusters (54% of responders *versus* 32% and 33%; p=0.02).

Conclusion 1-year treatment with lumacaftor–ivacaftor was associated with a significant visual improvement of bronchial disease on chest CT. Radiomics features on pre-therapeutic CT scans may help to predict lung function response under lumacaftor–ivacaftor.

Introduction

In recent years, new targeted therapies aimed at restoring the cystic fibrosis transmembrane conductance regulator (CFTR) channel function have been introduced and have dramatically changed the management of people with cystic fibrosis (CF) [1]. Among CFTR modulators, lumacaftor–ivacaftor was the first combination available for patients homozygous for the Phe508del mutation [2], who represent 40–50% of people with CF [3]. In phase III trials, lumacaftor–ivacaftor allowed a 1.9–3.7% improvement in per cent predicted forced expiratory volume in 1 s (FEV_1 % pred) and a 43% decrease in the rate of pulmonary exacerbations [4, 5].

Chest computed tomography (CT) is the gold standard for morphological assessment of lung diseases including CF [6]. Studies in people with CF have reported associations between lung function tests and/or rates of pulmonary exacerbations and morphological lung damage evaluated by CT scans [7, 8]. CT also provides additional information to spirometry, by showing new morphological damage in people with CF with stable FEV_1 % pred [9, 10]. However, chest CT has not been used as an end-point in trials evaluating the effects of CFTR modulators. Real-world observational studies have shown that ivacaftor induces marked improvement in airway wall thickening and extensive clearing of mucus plugging in patients with the Phe508del/G551D genotype [11, 12]. However, no data exist regarding CT monitoring of the effects of lumacaftor–ivacaftor in people with CF.

Despite resulting in improved lung function in the majority of patients, the response to treatment is variable from patient to patient, ranging from no or mild benefit to clinically significant improvement [13]. A previous study showed that the magnitude of response to lumacaftor–ivacaftor depended on FEV_1 % pred at treatment initiation [14], but the possible impact of CT morphological abnormalities on the response to lumacaftor–ivacaftor has not been assessed. However, patients with a comparable level of FEV_1 % pred can present with different disease phenotypes on chest CT, and only some CF-induced lesions are known to be reversible (*e.g.* under antibiotics). We therefore hypothesised that morphological phenotypes identified using machine learning-based methods on pre-therapeutic CT scans might predict the response to treatment.

The aims of this study were 1) to assess longitudinal morphological changes by comparing CT scans before and 1 year after initiation of lumacaftor–ivacaftor and 2) to use unsupervised machine learning on pre-therapeutic CT scans to identify structural abnormalities associated with lung function improvement under lumacaftor–ivacaftor.

Methods

Study population

This study was approved by the Institutional Review Board of the French Society for Respiratory Medicine (Société de Pneumologie de Langue Française, #2016-004), which waived the need for written consent. Data were extracted from a French real-world prospective multicentre observational study evaluating the safety and effectiveness of the lumacaftor–ivacaftor combination (NCT03475381) [15]. The study involved 845 patients with CF, including 292 adolescents aged ⩾12 years and 553 adults aged ⩾18 years, who were homozygous for the Phe508del *CFTR* mutation and were recruited from all 47 French CF reference centres. These patients had started lumacaftor–ivacaftor between January 1 and December 31, 2016. Available clinical data included demographic characteristics and FEV_1 % pred before initiation and after 12 months of treatment with lumacaftor–ivacaftor.

Patients were eligible for the first part of the study, which evaluated the effect of lumacaftor–ivacaftor on CT lung abnormalities, if there were available chest CT scans acquired 1) in stable condition (apart from a period of pulmonary exacerbation) within 12 months before initiation of lumacaftor–ivacaftor and 2) follow-up chest CT performed between 9 and 18 months after initiation of lumacaftor–ivacaftor. Patients who discontinued lumacaftor–ivacaftor within the first 12 months after initiation were excluded.

For the second part of the study, we aimed to use artificial intelligence to identify lung structural abnormalities before initiation of lumacaftor–ivacaftor that were associated with improvement in FEV_1 % pred after 1 year of lumacaftor–ivacaftor. To form the clustering dataset, we only selected patients with pre-therapeutic chest CT performed without contrast medium injection, with images reconstructed using a mediastinal kernel and a slice thickness ⩽2 mm.

Lung function response to lumacaftor–ivacaftor was defined as an increase in FEV_1 % pred ⩾5% after 12 months of lumacaftor–ivacaftor.

CT scans

Chest CT scans were acquired using 14 different 16–128 multi-slice CT models from five manufacturers. Acquisition parameters were specific to each centre with tube voltage ranging from 70 to 150 kV. Additional information regarding image acquisition and reconstruction is provided in the supplementary data. All images were fully anonymised, including the date of the CT examination.

Visual assessment of longitudinal changes in CT scans

All chest CT images were reviewed randomly using the same lung window setting (window width 1500 Hounsfield units, window level −450 Hounsfield units). Chest CT scans at treatment initiation and follow-up chest CT scans were analysed independently. Reviewers were blinded to the clinical information and the study date. The same radiologist with 4 years' experience (AC) scored all chest CT examinations using a modified Bhalla score [16]. This visual score rates the severity and the extent of bronchiectasis, peribronchial wall thickening, mucus plugging, generation of bronchial division involved, sacculations, bullae, emphysema, mosaic perfusion and collapse or consolidation. The total score ranges from 0 in the absence of morphological changes to 27. It also provides subscores for each anomaly; subscores for assessing bronchiectasis extent (range 0–3) and severity (range 0–3) were combined in a bronchiectasis subscore (range 0–6). To assess the intra- and inter-observer repeatability, 25 CT examinations were randomly selected. These CT examinations were again scored by the same observer in a second reading session 1 month apart and independently scored by a second observer (TNHT) with 3 years' experience in thoracic imaging.

Morphological phenotypes identification

This part of the study was conducted on the subset of patients from the clustering dataset. In a preprocessing step, all images were resampled to a 1-mm isometric voxel size using cubic interpolation. Lung segmentation, which consists of separating the lung from the chest wall and mediastinum, was then automatically performed on each axial slice using a publicly available lung segmentation model based on a three-dimensional deep-learning approach [17]. In case of segmentation error, the lung segmentation was manually edited. After lung segmentation, 120 radiomic features were extracted from the whole lung, including first-order statistics (19 features), shape descriptors (26 features) and textural features (75 features) (supplementary table e1). Feature extraction was performed using Python 3.6 (Python Software Foundation) and the Pyradiomics library [18]. The Pyradiomics library is a very popular library for radiomic studies that allows the extraction of 120 radiomic features per region of interest. These radiomic features, which are not perceptible to the human eye and allow a comprehensive description of the image, were used as the input in a k-means clustering method [19]. K-means clustering is a classical unsupervised machine learning method. The algorithm was run 200 times with random centroid initialisation, while the final clustering was chosen in terms of inertia. For our experiments, we used the Scikit-learn library [20]. Patients' characteristics at the time of initiation of lumacaftor–ivacaftor, Bhalla scores and the rate of response to treatment were compared between the clusters.

Statistical analysis

Statistical analysis was performed using the statistical software package R (v3.2.1, www.r-project.org). Groups were compared using a Wilcoxon test or an ANOVA for quantitative variables and a Chi-squared test for qualitative variables. Changes in the whole visual score and in each of its items, as well as changes in FEV_1 % pred between lumacaftor–ivacaftor initiation and follow-up, were compared using a Wilcoxon signed-rank test. A p-value <0.05 was considered statistically significant. The correlation between changes in the Bhalla scores and changes in FEV_1 % pred were evaluated with the Spearman R values and were interpreted as follows [21]: <0.2=very weak, 0.2–0.39=weak, 0.40–0.59=moderate, 0.60–0.79=strong and >0.8=very strong correlation. Intraclass correlation coefficients (ICCs) and Bland–Altman plots were used to assess intra- and inter-observer repeatability of visual scores. Excellent repeatability was assumed when ICC was ⩾0.8.

Results

Study population

Among the 845 patients participating in the French real-world lumacaftor–ivacaftor observational study, 476 had an available chest CT scan performed before initiation of lumacaftor–ivacaftor. Of these, 310 also had a follow-up chest CT scan. One patient who underwent an upper right lobectomy between the chest CT at lumacaftor–ivacaftor initiation and the follow-up CT scan was excluded. Another 26 patients who discontinued lumacaftor–ivacaftor prematurely (mostly owing to respiratory adverse events in the first

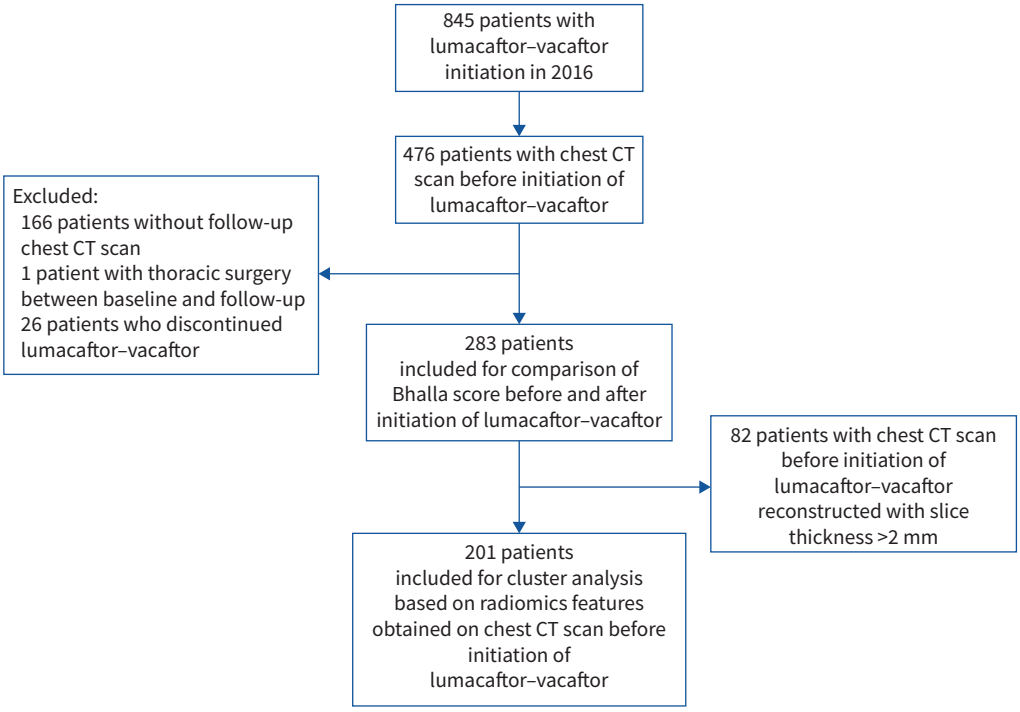

FIGURE 1 Flow-chart. CT: computed tomography.

weeks following initiation [15]) were excluded. This resulted in a cohort of 283 patients (figure 1) for analysing the effects of lumacaftor–ivacaftor on CT scan structural abnormalities. Among these 283 patients, 201 patients in whom the CT scan at lumacaftor–ivacaftor initiation was acquired without contrast and reconstructed with a soft kernel and a slice thickness ⩽2 mm were further included in the clustering dataset, aimed at identifying pre-therapeutic radiomics features associated with lung function improvement.

Patient characteristics at the time of initiation of lumacaftor–ivacaftor are presented in table 1. Patients were predominantly adults (n=174 of 283; 61%) and male (n=161 of 283, 57%). At lumacaftor–ivacaftor initiation, mean age was 23.4±9.1 years (range 12–52 years) and mean FEV_1 % pred was 65.9±19.6% (range 22–123%). The mean±SD increase in FEV_1 % pred at 1 year was 2.5±7.9% (p<0.001). 101 of 283 patients (36%) had an improvement in FEV_1 % pred ⩾5% in absolute value and were considered as responders. There were no significant differences in clinical characteristics between the whole 283 patient cohort and the 201 patient clustering subgroup (all p⩾0.14, table 1). The mean time interval between the pre-therapeutic chest CT scan and lumacaftor–ivacaftor initiation was 2.5±3.1 months (range 0.2–11.8 months). At lumacaftor–ivacaftor initiation, the mean Bhalla score was 12.7±3.7 points (range 1–22 points) (table 2). Its value moderately correlated with FEV_1 % pred at lumacaftor–ivacaftor initiation (R= −0.51, p<0.001) and number of intravenous antibiotic courses in the year prior to treatment initiation (R=0.52, p<0.001) (figure 2). The mean Bhalla score at lumacaftor–ivacaftor initiation was lower among responders than non-responders (12±3.5 points *versus* 13.1±3.7 points, p=0.01).

CT morphological changes under lumacaftor–ivacaftor

The mean time interval between lumacaftor–ivacaftor initiation and follow-up CT was 12.7±1.6 months (range 9.3–17.9 months). The mean difference between Bhalla score at lumacaftor–ivacaftor initiation and during follow-up was −1.40±1.53 points (range −8 to 2 points, p<0.001; table 2). Representative CT images depicting morphological changes occurring between initiation and 1 year after lumacaftor–ivacaftor initiation are shown in figure 3.

Among the subscores of the Bhalla score, the greatest improvement was observed for mucus plugging with a mean of difference of −0.58±0.88 points (range −2 to 2 points, p<0.001). Bronchial wall thickening and the severity of collapse or consolidations also significantly decreased under treatment with a mean difference of −0.35±0.62 points (range −3 to 2 points, p<0.001) and −0.24±0.52 points (range −2 to 2 points, p<0.001), respectively. There were no significant changes in severity and extent of bronchiectasis

TABLE 1 Patient characteristics at lumacaftor–ivacaftor initiation

	All	Clustering dataset	p-value[#]
Subjects, n	283	201	
Adults (⩾18 years)	174/283 (61%)	123/201 (61%)	0.95
Age, years	23.4±9.1	23.5±9.2	0.93
Male	161/283 (57%)	113/201 (56%)	0.88
Body mass index, kg·m^{-2}	19.9±2.3	19.9±2.3	0.95
Diabetes	79/283 (28%)	57/201 (28%)	0.92
Liver cirrhosis	10/150 (7%)	8/79 (10%)	0.36
FEV_1 % pred	65.9±19.6	68.4±19.6	0.21
Number of *i.v.* antibiotic courses in the previous 12 months	1.12±1.6	1.2±1.6	0.59
Airway pathogens			
Pseudomonas aeruginosa	170/283 (60%)	114/201 (57%)	0.46
Burkholderia cepacia	7/283 (2%)	6/201 (3%)	0.73
Methicillin-susceptible *Staphylococcus aureus*	193/283 (68%)	136/201 (68%)	0.90
Methicillin-resistant *S. aureus*	48/283 (17%)	37/201 (18%)	0.68
Haemophilus influenzae	40/283 (14%)	29/201 (14%)	0.93
Maintenance pulmonary medications at lumacaftor–ivacaftor initiation			
Azithromycin	178/283 (63%)	124/201 (62%)	0.79
Inhaled antibiotics	175/283 (62%)	123/201 (61%)	0.89
Dornase alpha	197/283 (70%)	144/201 (72%)	0.63
Inhaled bronchodilators	231/283 (82%)	168/201 (84%)	0.58
Inhaled corticosteroids	181/283 (64%)	127/201 (63%)	0.86
Oral corticosteroids	19/283 (7%)	8/201 (4%)	0.20
Lumacaftor–ivacaftor duration, days	374±40.4	347±91.1	0.14
Increase in FEV_1 % pred after lumacaftor–ivacaftor, %	2.5±7.9	2.6±9.1	0.63
Responders to lumacaftor–ivacaftor (increase in FEV_1 % pred ⩾5%)	101/283 (36%)	79/201 (39%)	0.42

Quantitative data are presented as mean±SD, qualitative data are presented as n (%). FEV_1: forced expiratory volume in 1 s; *i.v.*: intravenous. [#]: comparison assessed using Wilcoxon test for quantitative data and Chi-squared test for percentages.

(p=0.37), severity of sacculations/abscesses (p=0.85), severity of bullae (p=0.78) or severity of emphysema (p=1) (table 2, figure 4).

Visual CT score repeatability

There was excellent intra-observer agreement (ICC 0.93, 95% CI 0.85–0.97) and high inter-observer agreement (ICC 0.86, 95% CI 0.71–0.94) for visual CT scoring using the modified Bhalla score. Bland–Altman plots in figure 5 show that the 95% limits of agreement were −1.8 to 2.2 points for intra-observer agreement and −2.3 to 3.2 points for inter-observer agreement. The mean difference of the Bhalla score between chest CT scans acquired before lumacaftor–ivacaftor initiation and at follow-up (−1.40 points) was within this 95% limit of agreement. This means that the improvement was not sufficient to reliably capture the morphological improvement under lumacaftor–ivacaftor at the individual patient level.

CT morphological phenotype identification

Most lung segmentations did not require manual editing (n=123 of 201, 61%) and when done, manual edits were minor with a Dice similarity coefficient of 0.997±0.005 between the original and manually

TABLE 2 Morphological changes between lumacaftor–ivacaftor initiation and follow-up chest CT scans

Bhalla score	Chest CT at lumacaftor–ivacaftor initiation	Follow-up chest CT	p-value[#]
Total	12.7±3.7	11.4±3.7	<0.001
Mucus plugging	1.8±0.9	1.2±0.9	<0.001
Bronchial wall thickening	1.8±0.7	1.5±0.7	<0.001
Bronchiectasis	4.3±1.5	4.3±1.5	0.37
Bullae	0.1±0.3	0.1±0.3	0.78
Emphysema	0.2±0.6	0.2±0.5	1

Data are presented as mean±SD. CT: computed tomography. [#]: comparison assessed using the Wilcoxon paired test.

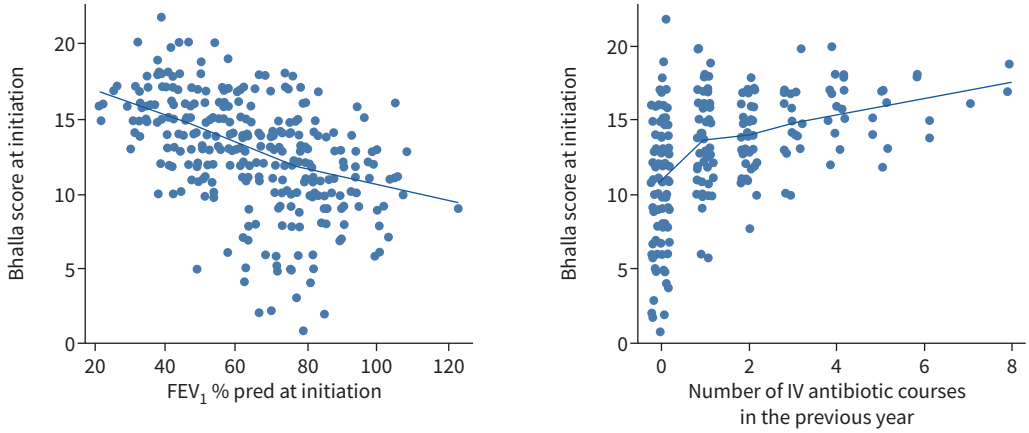

FIGURE 2 Correlation between Bhalla score and per cent predicted forced expiratory volume in 1 s (FEV_1) at initiation of lumacaftor–ivacaftor and number of intravenous (*i.v.*) antibiotic courses in the year prior to initiation of lumacaftor–ivacaftor.

edited segmentation. A clustering analysis based on the automated analysis of CT morphological characteristics (radiomics features) at lumacaftor–ivacaftor initiation led to the identification of three clusters of CF patients, based on the best split by k-means clustering. The number of clusters was selected according to the inflection point of the Calinski–Harabasz Index. Comparison of k-means clustering with other clustering algorithms (k-medians, density-based spatial clustering of applications with noise and

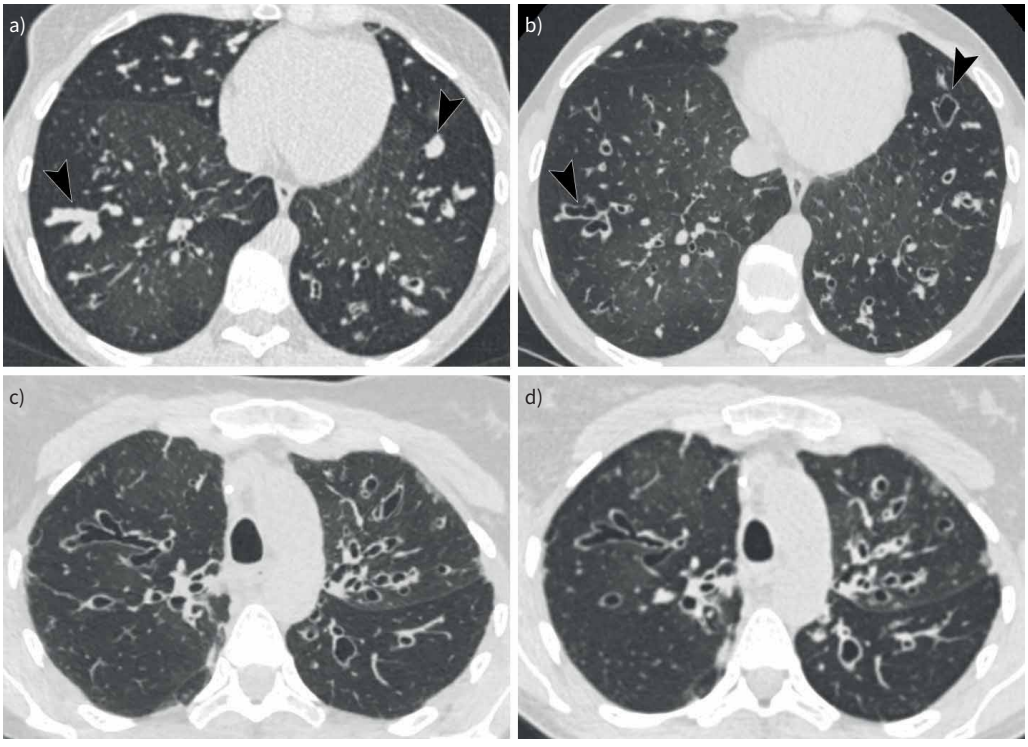

FIGURE 3 Side-by-side comparison of unenhanced chest computed tomography (CT) scans acquired 10 months before (a, c) and 1 year after (b, d) initiation of lumacaftor–ivacaftor in a 30-year-old female patient. Comparison shows the disappearance of mucus plugging (arrowheads) from image a to b, but a stability of the bronchiectasis from c to d. The Bhalla score was 16 before initiation of lumacaftor–ivacaftor and 13 after 1 year of lumacaftor–ivacaftor.

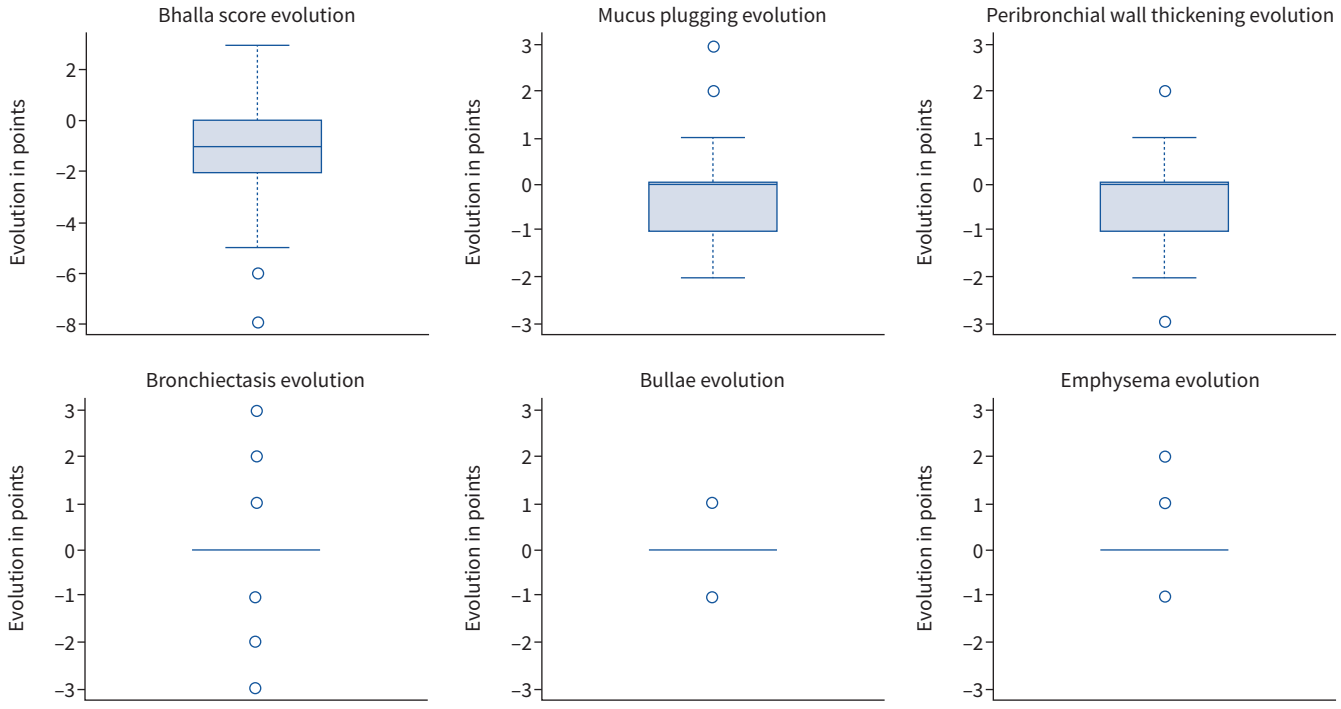

FIGURE 4 Boxplots showing the evolution of the different items of the Bhalla score (in points) between lumacaftor–ivacaftor initiation and follow-up chest computed tomography (CT) scans (n=283 patients). Central black line represents the median; boxplot represents the interquartile range; whiskers represent the 10th and 90th percentile; circles represent extreme values.

hierarchical clustering) using a silhouette coefficient and Calinski–Harabasz Index confirmed that k-means was the best clustering method for this analysis. The three clusters corresponding to the best split were of similar size but different clinical characteristics. These three clusters (A, B and C) comprised 80 patients (40%), 56 patients (28%) and 65 patients (32%) (total population n=201), respectively (table 3). Patients from clusters A and B had comparable demographic characteristics whereas patients from cluster C were significantly younger (19.6±8 years *versus* 25.6±10 years for cluster A and 24.9±9 years for cluster B, p<0.01). The proportion of patients with methicillin-resistant *Staphyloccocus aureus* colonisation was significantly lower in cluster C than in the other two clusters (6% for cluster C *versus* 26% for cluster A and 21% for cluster B, p=0.01). FEV_1 % pred was significantly higher in cluster C (75.7±17.8%) than in

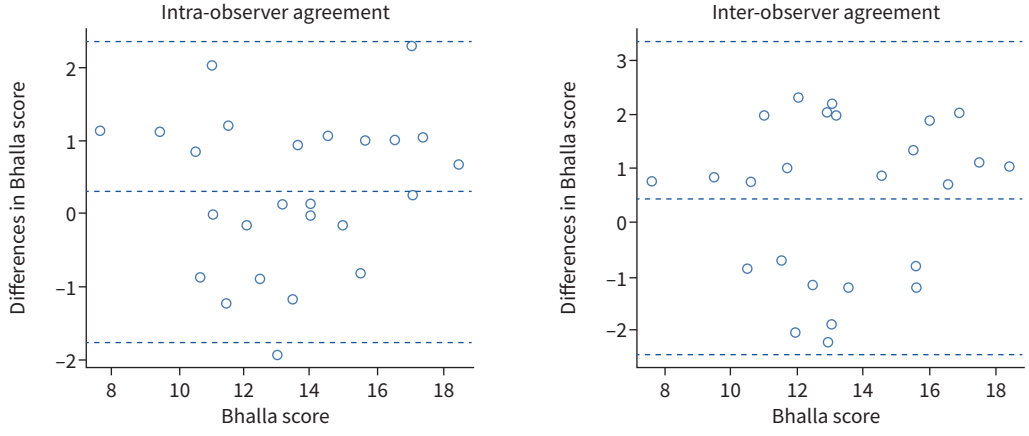

FIGURE 5 Bland–Altman plots showing the intra- and inter-observer agreements. Central dashed line represents the mean; inferior and superior dashed lines represet the 95% limit of agreement. Please note that jittering was used to view individual overlapping data points.

TABLE 3 Patients' clinical characteristics among the three clusters identified by k-means clustering of radiomics features

	Cluster A	Cluster B	Cluster C	p-value[#]
Subjects, n	80	56	65	
Responders to treatment (increase in FEV_1 % pred ⩾5%)	26 (32%)	18 (33%)	35 (54%)	0.02
Age at treatment initiation, years	25.6±10	24.9±9	19.6±8	<0.001
Adults (⩾18 years)	56 (70%)	40 (71%)	27 (42%)	<0.001
Male	43 (54%)	38 (68%)	32 (49%)	0.10
FEV_1 % pred	60.5±17.3	71±20.8	75.7±17.8	<0.001
Airway pathogens				
Pseudomonas aeruginosa	49(61%)	35 (62%)	30 (46%)	0.11
Burkholderia cepacia	2 (2%)	2 (4%)	2 (3%)	0.94
Methicillin-susceptible *Staphylococcus aureus*	50 (62%)	36 (64%)	50 (77%)	0.15
Methicillin-resistant *S. aureus*	21 (26%)	12 (21%)	4 (6%)	0.01
Haemophilus influenzae	10 (12%)	11 (20%)	8 (12%)	0.43
Maintenance pulmonary medications at lumacaftor–ivacaftor initiation				
Azithromycin	51 (64%)	30 (54%)	43 (66%)	0.32
Inhaled antibiotics	52 (65%)	37 (66%)	34 (52%)	0.20
Dornase alpha	50 (62%)	38 (68%)	56 (86%)	0.01
Inhaled bronchodilators	68 (85%)	45 (80%)	55 (85%)	0.74
Inhaled corticosteroids	49 (61%)	32 (57%)	46 (71%)	0.27
Oral corticosteroids	2 (2%)	2 (4%)	4 (6%)	0.53
Bhalla score at lumacaftor–ivacaftor initiation				
Total score	13.8±3.8	12.2±3.1	10.5±3.7	<0.001
Mucus plugging	2.1±0.8	1.6±0.8	1.3±0.9	<0.001
Peribronchial wall thickening	2.1±0.7	1.7±0.7	1.5±0.7	<0.001
Bronchiectasis[¶]	4.3±1.7	3.9±1.3	3.5±1.6	<0.001
Bullae	0.1±0.6	0.1±0.4	0.1±0.4	0.94
Emphysema	0.2±0.4	0.3±0.4	0.2±0.4	0.19

Quantitative data are presented as mean±SD, qualitative data are presented as n (%). FEV_1: forced expiratory volume in 1 s. [#]: comparison between the three clusters assessed using ANOVA for quantitative data and Chi-squared test for quantitative data. [¶]: score from 0 to 6 (see Methods).

cluster A (60.5±17.3%) and B (71±20.8%) (p<0.001). CT visual analysis also showed a decreasing severity gradient from cluster A (total Bhalla score 13.8±3.8 points) to B (total Bhalla score 12.2±3.1 points) to C (total Bhalla score 10.5±3.7 points) (p<0.001). Patients from cluster C had less mucus plugging (1.3±0.9 points *versus* 2.1±0.8 points in cluster A and 1.6±0.8 points in cluster B, p<0.01), less peribronchial wall thickening (1.5±0.7 points *versus* 2.1±0.7 points in cluster A and 1.7±0.7 points in cluster B, p<0.01) and less bronchiectasis (3.5±1.6 points *versus* 4.3±1.7 points in cluster A and 3.9±1.3 points in cluster B, p<0.01). Interestingly, patients from cluster C, who had less morphologically and functionally severe disease, were better responders to treatment than the others (54% of responders in cluster C *versus* 32% in cluster A and 33% in cluster B, p=0.02).

Discussion

We found that the improvement in lung function following initiation of lumacaftor–ivacaftor in people with CF was associated with significant improvement of bronchial abnormalities on chest CT based on visual scoring. We were also able to identify a subgroup of patients with reduced lung structural abnormalities at lumacaftor–ivacaftor initiation, which predicted higher rates of FEV_1 % pred response to lumacaftor–ivacaftor.

Using data from a large nationwide real-world observational study, we showed that treatment with lumacaftor–ivacaftor was associated with morphological improvement on CT, which was related to decreased mucus plugging and peribronchial wall thickening. These lesions are known to be potentially reversible, as shown by studies evaluating the effects of antibiotics on CT scans [10, 22]. Conversely, bronchiectasis, bullae, emphysema and sacculations did not significantly improve under treatment with lumacaftor–ivacaftor. These results are in line with previous reports on the morphological improvement in CF patients with gating mutations treated by ivacaftor [11, 12]. In these studies, morphological improvement was mainly related to decreased mucus plugging and peribronchial thickening [12, 23, 24]. There are conflicting data regarding the evolution of bronchiectasis under ivacaftor. Some authors have

reported that treatment with ivacaftor is associated with a significant decrease in bronchiectasis score [12] whereas others have reported an increase, which could be due to the natural evolution of the disease [24]. The stability of the bronchiectasis score after 1 year in our study is consistent with the fact that bronchiectasis results from irreversible changes due to the destruction of elastic tissues [25].

We found a significant correlation between visual CT scores of disease severity and FEV_1 % pred ($R = -0.51$). This correlation was in the range of those previously reported for the Bhalla score and other visual scores ($R = -0.52$ to -0.85) [12, 24, 26]. However, we did not observe a significant correlation between morphological improvement and FEV_1 % pred improvement, whereas such correlation has previously been reported in patients with gating mutations treated with ivacaftor [24]. We speculate that this difference could be related to a reduced magnitude of effect on lung function in patients homozygous for the Phe508del mutation treated with lumacaftor–ivacaftor, as compared to the effects of ivacaftor in patients with gating mutations.

The improvement in the Bhalla score, which quantifies morphological changes, was relatively small yet significant, with a mean decrease of only 1.40 points out of 27. This value was within the 95% limit of agreement of −2.3 to 3.2 for inter-observer repeatability. This suggests that the Bhalla score can be used for assessing structural changes at the cohort level, but cannot be reliably used to capture the morphological improvement under lumacaftor–ivacaftor at the individual patient level. Of course, this conclusion could be different when assessing the effects of newer modulator combinations (*e.g.* the recently released elexacaftor–tezacaftor–ivacaftor), which show more potent effects on lung function improvement [27–29].

In the present study, machine learning was used to identify different structural abnormalities observed on chest CT prior to initiation of lumacaftor–ivacaftor that would predict lung function response under lumacaftor–ivacaftor. Previous attempts to describe lung disease phenotypes in CF patients included the development of the Severe Advanced Lung Disease (SALD) scoring system [9]. The SALD scoring system is based on visual quantification of airway disease using grid cells applied to the images. Using this scoring system in severe pre-transplantation CF patients, Loeve *et al*. [9] showed that at one end of the spectrum, patients had predominantly infection/inflammation-related changes and at the other end predominantly air trapping/hypoperfusion. We used a data-driven approach entirely based on automated image analysis. An unsupervised machine learning method combined with radiomics feature extraction allowed us to identify three different morphological clusters on chest CT. We identified a cluster (cluster C) of better responders who were younger and had less morphologically and functionally severe disease. Interestingly, although mucoid impactions and bronchial thickening were reversible lesions under treatment, they were less prevalent in the phenotype better responding to treatment. These findings further support the need for introduction of CFTR modulators earlier in life.

Our study has several limitations. First, owing to its observational design, CT scans at lumacaftor–ivacaftor initiation and during follow-up were not available for all patients, because the decision to perform CT was left to the treating physicians. Second, the use of other visual scores, such as Brody's score and its derivatives, could have allowed a finer quantification of morphological changes by assessing the disease on a segmental instead of lobar level. However, most of the scored items are similar between the methods and we preferred Bhalla's score because of its shorter time for measurement, which was important because of the rather large cohort. The Bhalla score allowed us to show significant improvement under treatment. Last, we only used unsupervised machine learning methods for clustering. Indeed, despite this cohort being the largest of people with CF treated by lumacaftor–ivacaftor, we were not able to use deep learning to predict response to treatment owing to the relatively limited size of the cohort.

In conclusion, we found that 1 year of treatment by lumacaftor–ivacaftor in people with CF homozygous for the Phe508del mutation resulted in significant improvement of morphological changes on CT, characterised by decreased mucus plugging and airway wall thickening. In addition, we identified a cluster of patients with milder structural abnormalities on CT scan prior to initiation of lumacaftor–ivacaftor. This cluster was associated with a markedly higher probability of lung function improvement under lumacaftor–ivacaftor. The better response rate observed in patients with younger age and less severe structural abnormalities concurs with the concept that CFTR modulators should be started early in life. Analyses of CT scans using visual scoring and artificial intelligence could prove important in assessing response to therapy, including newer CFTR modulator combinations, in people with CF.

This article has been revised according to the correction published in the March 2023 issue of the *European Respiratory Journal*.

Acknowledgements: The authors thank URC-CIC Paris Descartes Necker Cochin (Caroline Tourte, Guillaume Masson) for the implementation of the study. The authors also thank the French CF Registry team (Lydie Lemonnier, Clémence Dehilotte) and Jean-Louis Paillasseur (Effi-Stat, Paris, France) for help with data management of the study.

Participating investigators in the French CF reference network study group: Julie Mounard, Claire Poulet, Cinthia Rames (Amiens); Christine Person, Françoise Troussier, Thierry Urban (Angers); Marie-Laure Dalphin, Jean-Claude Dalphin, Didier Pernet, Bénédicte Richaud-Thiriez (Besançon); Stéphanie Bui, Mickael Fayon, Julie Macey-Caro (Bordeaux); Karine Campbell, Muriel Laurans (Caen); Corinne Borderon, Marie-Christine Heraud, André Labbé, Sylvie Montcouquiol (Clermont-Ferrand); Laurence Bassinet, Natascha Remus (Créteil); Annlyse Fanton, Anne Houzel-Charavel, Frédéric Huet, Stéphanie Perez-Martin (Dijon); Amale Boldron-Ghaddar, Manuela Scalbert (Dunkerque); Laurent Mely (Giens); Boubou Camara, Catherine Llerena, Isabelle Pin, Sébastien Quétant (Grenoble); Aurélie Cottereau, Antoine Deschildre, Alice Gicquello, Thierry Perez, Lidwine Stervinou-Wemeau, Caroline Thumerelle, Benoit Wallaert, Nathalie Wizla (Lille); Jane Languepin, Céline Ménétrey, Magalie Dupuy-Grasset (Limoges); Lucie Bazus, Clelia Buchs, Virginie Jubin, Marie-Christine Werck-Gallois, Catherine Mainguy, Thomas Perrin, Philippe Reix, Agnès Toutain-Rigolet (Lyon Pédiatrie); Isabelle Durieu, Stéphane Durupt, Quitterie Reynaud, Raphaele Nove-Josserand (Lyon adultes); Melisande Baravalle-Einaudi, Bérangère Coltey, Nadine Dufeu, Jean-Christophe Dubus, Nathalie Stremler (Marseille); Davide Caimmi, Raphaël Chiron (Montpellier); Yves Billon, Jocelyne Derelle, Sébastien Kieffer, Anne-Sophie Pichon, Cyril Schweitzer, Aurélie Tatopoulos (Nancy); Sarah Abbes, Tiphaine Bihouée, Isabelle Danner-Boucher, Valérie David, Alain Haloun, Adrien Tissot (Nantes); Sylvie Leroy, Carole Bailly-Piccini (Nice); Annick Clément, Harriet Corvol, Aline Tamalet (Paris, Trousseau); Pierre-Régis Burgel, Isabelle Honoré, Dominique Hubert, Reem Kanaan, Clémence Martin (Paris, Cochin); Cécile Bailly, Frédérique Chédevergne, Jacques De Blic, Brigitte Fauroux, Murielle Le Bourgeois, Isabelle Sermet-Gaudelus (Paris, Necker); Bertrand Delaisi, Michèle Gérardin, Anne Munck (Paris, Robert Debré); Michel Abély, Bruno Ravoninjatovo (Reims); Chantal Belleguic, Benoit Desrues, Graziella Brinchault (Rennes); Michel Dagorne, Eric Deneuville, Sylvaine Lefeuvre (Rennes-Saint Brieuc); Anne Dirou, Jean Le Bihan, Sophie Ramel (Roscoff); Stéphane Dominique, Christophe Marguet (Rouen); Annabelle Payet (La Réunion); Romain Kessler, Michele Porzio, Vincent Rosner, Laurence Weiss (Strasbourg); Sandra de Miranda, Dominique Grenet, Abdoul Hamid, Clément Picard (Suresnes); François Brémont, Alain Didier, Géraldine Labouret, Marie Mittaine, Marlène Murris-Espin, Laurent Têtu (Toulouse); Laure Cosson, Charlotte Giraut, Anne-Cécile Henriet, Julie Mankikian, Sophie Marchand (Tours); Sandrine Hugé, Véronique Storni (Vannes); Emmanuelle Coirier-Duet (Versailles).

Conflict of interest: A. Campredon has nothing to disclose. E. Battistella has nothing to disclose. C. Martin reports lecture payments or honoraria from Chiesi and Zambon, outside the submitted work. I. Durieu has nothing to disclose. L. Mely has nothing to disclose. C. Marguet reports consulting fees from Gleamer; lecture payments or honoraria from Vertex, Viatis and Zambon; support for attending meetings and/or travel from Zambon; and participation on a Data Safety Monitoring Board or Advisory Board for Zambon and Viatis; outside the submitted work. C. Belleguic has nothing to disclose. M. Murris-Espin has nothing to disclose. R. Chiron has nothing to disclose. A. Fanton has nothing to disclose. S. Bui reports payment for expert testimony for inhaled antibiotics for Zambon, outside the submitted work. M. Reynaud-Gaubert has nothing to disclose. P. Reix has nothing to disclose. T-N. Hoang-Thi has nothing to disclose. M. Vakalopoulou has nothing to disclose. M-P. Revel has nothing to disclose. J. Da Silva has nothing to disclose. P-R. Burgel reports grants or contracts from Vertex and GSK; consulting fees from AstraZeneca, Chiesi, GSK, Insmed, Vertex and Zambon; and lecture payments or honoraria from Pfizer and Novartis; outside the submitted work. G. Chassagnon reports lecture payments or honoraria from Chiesi, outside the submitted work.

Support statement: This work was funded by grants from Vaincre la Mucoviscidose, Société Française de la Mucoviscidose and Legs Pascal Bonnet. Funding sources were not involved in the study design; the collection, analysis and interpretation of the data; the writing of the report; and the decision to submit the article for publication. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Bell SC, Mall MA, Gutierrez H, *et al.* The future of cystic fibrosis care: a global perspective. *Lancet Respir Med* 2020; 8: 65–124.
2 Wainwright CE, Elborn JS, Ramsey BW, *et al.* Lumacaftor–ivacaftor in patients with cystic fibrosis homozygous for Phe508del CFTR. *N Engl J Med* 2015; 373: 220–231.
3 Veit G, Avramescu RG, Chiang AN, *et al.* From CFTR biology toward combinatorial pharmacotherapy: expanded classification of cystic fibrosis mutations. *Mol Biol Cell* 2016; 27: 424–433.
4 Elborn JS, Ramsey BW, Boyle MP, *et al.* Efficacy and safety of lumacaftor/ivacaftor combination therapy in patients with cystic fibrosis homozygous for Phe508del *CFTR* by pulmonary function subgroup. *Lancet Respir Med* 2016; 4: 617–626.

5 Konstan MW, McKone EF, Moss RB, *et al.* Assessment of safety and efficacy of long-term treatment with combination lumacaftor and ivacaftor therapy in patients with cystic fibrosis homozygous for the F508del-*CFTR* mutation (PROGRESS): a phase 3, extension study. *Lancet Respir Med* 2017; 5: 107–118.
6 Kołodziej M, de Veer MJ, Cholewa M, *et al.* Lung function imaging methods in cystic fibrosis pulmonary disease. *Respir Res* 2017; 18: 96.
7 Brody AS, Klein JS, Molina PL, *et al.* High-resolution computed tomography in young patients with cystic fibrosis: distribution of abnormalities and correlation with pulmonary function tests. *J Pediatr* 2004; 145: 32–38.
8 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: 1128–1132.
9 Loeve M, Krestin GP, Rosenfeld M, *et al.* Chest computed tomography: a validated surrogate endpoint of cystic fibrosis lung disease? *Eur Respir J* 2013; 42: 844–857.
10 Horsley AR, Davies JC, Gray RD, *et al.* Changes in physiological, functional and structural markers of cystic fibrosis lung disease with treatment of a pulmonary exacerbation. *Thorax* 2013; 68: 532–539.
11 Hoare S, McEvoy S, McCarthy CJ, *et al.* Ivacaftor imaging response in cystic fibrosis. *Am J Respir Crit Care Med* 2014; 189: 484.
12 Sheikh SI, Long FR, McCoy KS, *et al.* Computed tomography correlates with improvement with ivacaftor in cystic fibrosis patients with G551D mutation. *J Cyst Fibros* 2015; 14: 84–89.
13 Jones AM. Royal Society of Medicine cystic fibrosis symposium 2017. *Paediatr Respir Rev* 2018; 27: 1.
14 Burgel P-R, Durieu I, Chiron R, *et al.* Clinical response to lumacaftor–ivacaftor in patients with cystic fibrosis according to baseline lung function. *J Cyst Fibros* 2021; 20: 220–227.
15 Burgel P-R, Munck A, Durieu I, *et al.* Real-life safety and effectiveness of lumacaftor–ivacaftor in patients with cystic fibrosis. *Am J Respir Crit Care Med* 2020; 201: 188–197.
16 Helbich TH, Heinz-Peer G, Fleischmann D, *et al.* Evolution of CT findings in patients with cystic fibrosis. *AJR Am J Roentgenol* 1999; 173: 81–88.
17 Hofmanninger J, Prayer F, Pan J, *et al.* Automatic lung segmentation in routine imaging is a data diversity problem, not a methodology problem. *ArXiv* 2020; preprint [https://doi.org/10.1186/s41747-020-00173-2].
18 van Griethuysen JJ, Fedorov A, Parmar C, *et al.* Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017; 77: e104–e107.
19 MacQueen J. Some methods for classification and analysis of multi-variate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1967; 5: 281–297.
20 Pedregosa F, Varoquaux G. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011; 12: 2825–2830.
21 Evans JD. Straightforward Statistics for the Behavioral Sciences. Pacific Grove, CA, Thomson Brooks/Cole Publishing Co., 1996.
22 Shah RM, Sexauer W, Ostrum BJ, *et al.* High-resolution CT in the acute exacerbation of cystic fibrosis: evaluation of acute findings, reversibility of those findings, and clinical correlation. *AJR Am J Roentgenol* 1997; 169: 375–380.
23 Muhlebach M, Clancy J, Heltshe S, *et al.* Biomarkers for cystic fibrosis drug development. *J Cyst Fibros* 2016; 15: 714–723.
24 Chassagnon G, Hubert D, Fajac I, *et al.* Long-term computed tomographic changes in cystic fibrosis patients treated with ivacaftor. *Eur Respir J* 2016; 48: 249–252.
25 Whitwell F. A study of the pathology and pathogenesis of bronchiectasis. *Thorax* 1952; 7: 213–239.
26 Cademartiri F, Luccichenti G, Palumbo AA, *et al.* Predictive value of chest CT in patients with cystic fibrosis: a single-center 10-year experience. *AJR Am J Roentgenol* 2008; 190: 1475–1480.
27 Middleton PG, Mall MA, Dřevínek P, *et al.* Elexacaftor–tezacaftor–ivacaftor for cystic fibrosis with a single Phe508del allele. *N Engl J Med* 2019; 381: 1809–1819.
28 Heijerman HGM, McKone EF, Downey DG, *et al.* Efficacy and safety of the elexacaftor plus tezacaftor plus ivacaftor combination regimen in people with cystic fibrosis homozygous for the F508del mutation: a double-blind, randomised, phase 3 trial. *Lancet* 2019; 394: 1940–1948.
29 Burgel P-R, Durieu I, Chiron R, *et al.* Rapid improvement after starting elexacaftor-tezacaftor-ivacaftor in patients with cystic fibrosis and advanced pulmonary disease. *Am J Respir Crit Care Med* 2021; 204: 64–73.