



**HAL**  
open science

## **Artificial intelligence-driven volumetric CT outcome score in cystic fibrosis: longitudinal and multicenter validation with/without modulators treatment**

Amel Imene Hadj Bouzid, Stephanie Bui, Ilyes Benlala, Patrick Berger, Antoine Hutt, Renan Liberge, Paul Habert, Jean-Yves Gaubert, Marie Baque-Juston, Baptiste Morel, et al.

### ► To cite this version:

Amel Imene Hadj Bouzid, Stephanie Bui, Ilyes Benlala, Patrick Berger, Antoine Hutt, et al.. Artificial intelligence-driven volumetric CT outcome score in cystic fibrosis: longitudinal and multicenter validation with/without modulators treatment. *European Radiology*, 2024, 35, pp.815-827. <10.1007/s00330-024-11019-5>. <hal-04902462>

**HAL Id: hal-04902462**

**<https://hal.science/hal-04902462v1>**

Submitted on 20 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## **Artificial intelligence-driven volumetric CT outcome score in cystic fibrosis: longitudinal and multicenter validation with/without modulators treatment**

Amel Imene Hadj Bouzid 1, Stephanie Bui 1 2 3, Ilyes Benlala 1 2 3, Patrick Berger 1 2 3, Antoine Hutt 4, Renan Liberge 5, Paul Habert 6, Jean-Yves Gaubert 6, Marie Baque-Juston 7, Baptiste Morel 8, Gilbert Ferretti 9, Baudouin Denis de Senneville 10, François Laurent 1 2 3, Julie Macey 1 2 3, Gaël Dournes 11 12 13

1 Univ. Bordeaux, , INSERM, Centre de Recherche Cardio-Thoracique de Bordeaux, U1045, CIC 1401, F-33600, Pessac, France.

2 CHU Bordeaux, Service d'Imagerie Thoracique et Cardiovasculaire, Service des Maladies Respiratoires, Service d'Exploration Fonctionnelle Respiratoire, Paediatric Cystic Fibrosis Reference Center (CRCM), CIC 1401, F-33600, Pessac, France.

3 INSERM, U1045, Centre de Recherche Cardio-Thoracique de Bordeaux, CIC 1401, F-33600, Pessac, France.

4 Department of Thoracic Imaging, Heart & Lung Institute, Lille, Cedex, France.

5 Department of Radiology, CHU Nantes, F-44000, Nantes, France.

6 Imaging Department, Hopital La Timone, APHM, Aix Marseille University, Marseille, France.

7 Paediatric Radiology Department, Hôpitaux Pédiatriques de Nice CHU-Lenval, Nice, France.

8 Pediatric Radiology Department, Clocheville Hospital, CHRU de Tours, Tours, France.

9 Department of Radiology, Grenoble-Alpes University Hospital, 38000, Grenoble, France.

10 CNRS, Bordeaux INP, LaBRI, UMR 5800, Bordeaux INP, UMR 5251, F-33400, Talence, France.

11 Univ. Bordeaux, INSERM, Centre de Recherche Cardio-Thoracique de Bordeaux, U1045, CIC 1401, F-33600, Pessac, France. [gael.dournes@chu-bordeaux.fr](mailto:gael.dournes@chu-bordeaux.fr).

12 CHU Bordeaux, Service d'Imagerie Thoracique et Cardiovasculaire, Service des Maladies Respiratoires, Service d'Exploration Fonctionnelle Respiratoire, Paediatric Cystic Fibrosis Reference Center (CRCM), CIC 1401, F-33600, Pessac, France. [gael.dournes@chu-bordeaux.fr](mailto:gael.dournes@chu-bordeaux.fr).

13 INSERM, U1045, Centre de Recherche Cardio-Thoracique de Bordeaux, CIC 1401, F-33600, Pessac, France. [gael.dournes@chu-bordeaux.fr](mailto:gael.dournes@chu-bordeaux.fr).

## **Abstract**

**Objectives.** Holistic segmentation of CT structural alterations with 3D deep learning has recently been described in cystic fibrosis (CF), allowing the measurement of Normalized Volumes of Airway Abnormalities (NOVAA-CT) as an automated quantitative outcome. Clinical validations are needed, including longitudinal and multicenter evaluations.

**Material and Methods.** The validation study was retrospective between 2010 and 2023. CF patients undergoing Elexacaftor/Tezacaftor/Ivacaftor (ETI) or corticosteroids for allergic broncho-pulmonary aspergillosis (ABPA) composed the monocenter ETI and ABPA groups, respectively. Patients from six geographically-distinct Institutions composed a multicenter External group. All patients had completed CT and pulmonary function test (PFT), with a second assessment at one year in case of ETI or ABPA treatment. NOVAA-CT quantified bronchiectasis, peribronchial thickening, bronchial mucus, bronchiolar mucus, collapse/consolidation volumes, and Total Abnormal Volume (TAV) as their sum. Two observers evaluated the visual Bhalla score.

**Results.** 139 CF patients (median age, 15 years [interquartile range: 13-25]) were evaluated. All correlations between NOVAA-CT to both PFT and Bhalla score were significant in the ETI (n=60), ABPA (n=20), and External groups (n=59), such as the normalized TAV ( $\rho \geq 0.76$ ;  $p < 0.001$ ). In both ETI and ABPA groups, there were significant longitudinal improvements in peribronchial thickening, bronchial mucus, bronchiolar mucus and collapse/consolidation ( $p \leq 0.001$ ). An additional reversibility in bronchiectasis volume was quantified with ETI ( $p < 0.001$ ). Intraclass correlation coefficient of reproducibility was  $> 0.99$ .

**Conclusion.** NOVAA-CT automated scoring demonstrates validity, reliability and responsiveness for monitoring CF severity over an entire lung, and quantifies therapeutic effects on lung structure at CT, such as the volumetric reversibility of airway abnormalities with ETI.

**Clinical relevance statement:** Visual scorings are subjective, lack sensitivity and reproducibility to assess bronchial changes in cystic fibrosis. NOVAA-CT automated 3D outcome enables reproducible and holistic monitoring of CF severity over an entire lung for objective management and quantitative endpoint during therapeutic trials.

## **Keywords**

CT; Artificial intelligence; Deep learning; Cystic fibrosis; Airways

## **Key points**

1. Visual scoring methods lack sensitivity and reproducibility to assess longitudinal bronchial changes in cystic fibrosis.
2. Artificial intelligence-driven volumetric CT scoring correlates longitudinally to the disease severity in both monocenter and multicenter settings ( $p < 0.001$ ), and reliably improves with Elexacaftor/Tezacaftor/Ivacaftor or corticosteroid treatments ( $p < 0.01$ ).
3. Artificial intelligence-driven volumetric CT scoring enables reproducible monitoring the lung disease severity in CF and quantifies longitudinal therapeutic effects on lung structure, such as the volumetric reversibility of bronchial alterations with Elexacaftor/Tezacaftor/Ivacaftor.

## INTRODUCTION

Cystic fibrosis (CF) is one of the most frequent genetic diseases in Caucasians, affecting between 1 in every 6000 and 1 in every 3000 children in Europe and North America[1]. The advent of effective treatments has revolutionized the prognosis of CF-related lung disease[2, 3], and objective tools for reproducibly monitoring the longitudinal evolution of individuals with CF have become crucial. Computed Tomography (CT) is the standard of reference for assessing bronchial and parenchymal alterations of the lung structure *in vivo*[4]. To use it as an outcome measure[5], several visual scoring systems have been proposed[6–9]. However, visual scoring systems are subjective, time-consuming, and may lack reproducibility[10]. In addition, for practical handling use, the complex information available on CT is simplified, at the expense of information loss. For instance, up to two lung lobes may be required to allow for a single point of variation using the Bhalla scoring system[7]. Thus, consensus on their clinical use is still lacking[10].

To cope with these known shortcomings, quantitative scorings have been reported, without[11–14], or with artificial intelligence (AI) algorithms with deep learning[15–17]. Interestingly, AI allows for innovative computer vision tasks[18] to reach specific semantic labeling of five structural hallmarks of CF, *i.e.* bronchiectasis, peribronchial thickening, bronchial mucus, bronchiolar mucus, and collapse/consolidation[15]. Besides, AI can reach holistic segmentations, meaning that the automated segmentations are made at the voxel level over an entire CT scan. Conversely, human readers would spend several hours or days performing this task manually. Recently, three dimensional (3D)-convolutional neural networks (CNN) was found more effective than two dimensional (2D)-CNN to perform this multilabeling process, with high accuracy and specificity[17]. These 3D segmentations can be translated in the form of volumes of lesions, which may offer an opportunity for an automated quantitative scoring system. In this manuscript, we will introduce it as the Normalized Volume of Airway Abnormalities CT score (NOVAA-CT).

Nevertheless, previous descriptions of AI-driven segmentations related mainly to the semantic performances of the algorithms, demonstrating precise similarities as compared to experts' manual ground truths[15, 17]. However, to consider it as an outcome measure, clinical evaluation of validity, reliability and longitudinal responsiveness are required[19]. Indeed, a valid surrogate endpoint needs to clinically reflect the disease severity, correlate to established outcomes, and improve after treatment[5]. In addition, because NOVAA-CT is an AI-based application, evaluations with temporally-distinct and geographically-distinct tests are suitable to assess the generalizability[18].

Therefore, we hypothesized that NOVAA-CT could be a valid, reliable and responsive outcome to quantify airway severity changes in CF, with possible generalizability. Thus, the study aimed at assessing the correlations between NOVAA-CT and two established surrogate endpoints in CF, *i.e.* the Forced Expiratory Volume in 1-second percentage predicted (FEV1%p) and the CT Bhalla score,

including temporally distinct and external evaluations. Secondary objectives were (ii) to assess the longitudinal modifications of NOVAA-CT in CF patients undergoing Elexacaftor/Tezacaftor/Ivacaftor (ETI) or (ii) corticosteroid treatments for allergic broncho-pulmonary aspergillosis (ABPA), and (iii) to assess the reproducibility and acquisition time of the method.

## **MATERIAL AND METHODS**

### *Study populations*

The retrospective study involved three independent cohorts across seven CF reference centers, and all patients provided written informed consent for reusing data from their medical records after Institutional Review Board (IRB) approval (study registration number NCT04760548; IRB registration number CRM-2312-384).

Firstly, consecutive CF patients undergoing ETI (ETI group) were recruited between 2021 and 2023 in two reference centers dedicated to children and adults, respectively, at a single Institution. Inclusion criteria were an age greater than 12 years, a diagnosis of CF made on sweat chloride and/or genetic test, a heterozygous or homozygous deltaF508 mutation, a non-contrast-enhanced CT scan and pulmonary function tests (PFT) performed during a routine annual visit at two timepoints, before (M0) and 1 year (M12) after initiation of ETI treatment (Kaftrio®, Vertex Pharmaceutical) in a period of disease stability.

Secondly, a temporally-distinct group of CF patients undergoing corticosteroids for ABPA (ABPA group) was recruited from the same centers between 2010 and 2014. Inclusion criteria were the same as for the ETI group, except for the age superior to 6 years and the need for ABPA treatment[20] instead of ETI. Diagnosis of ABPA was made according to international consensus criteria[21] at baseline and a second CT scan had to be available 1 year after treatment, in a period of disease stability.

Thirdly, a geographically-distinct group was recruited (External group). For this, six external Institutions conferred their agreement for a retrospective ancillary analysis of CT and PFT data collected between 2018 and 2022 during a trial (parent study number NCT03357562). Inclusion criteria were an age greater than 8 years old, a diagnosis of CF made on sweat chloride and/or genetic test, a non-contrast-enhanced CT scan, and pulmonary function tests (PFT) performed during a routine annual visit.

In all three ETI, ABPA, and External groups (Figure 1), disease management was performed according to a standard of care[20].

### *Pulmonary function test*

Body plethysmography was used to assess lung volumes. Functional parameters were recorded in liters and expressed as percentages of predicted values: FEV<sub>1</sub>, forced vital capacity (FVC), total lung capacity (TLC), and residual volume (RV). European Respiratory Society and American Thoracic Society guidelines were chosen for reference values[22, 23].

### *CT scans*

Details on CT scans are provided in e-Table 1. Briefly, chest CT scans were done on four multidetector CT scanners from two major manufacturers (Siemens Healthcare, Erlangen, Germany; GE HealthCare, USA). Each independent group had distinct CT scan characteristics; however, CT protocols were standardized in patients who underwent repeat imaging during longitudinal evaluation (e-Table1). The cooperative young and adult patients were instructed before and during the procedure to hold their breath at the end inspiration. The effective exam dose ranged from 1.2 to 1.5 mSv. Voxel size ranged from 0.625 to 1.25 mm<sup>3</sup>.

### *Automated quantitative measurement of NOVAA-CT*

NOVAA-CT was implemented according to a previously validated multilabeling method[15] which was implemented into a 3D architecture using the 3D-nnUnet algorithm[17], available at <https://github.com/MIC-DKFZ/nnUNet> (e-Method 1-2-3). Herein, additional training was performed on 250 000 training samples (including data augmentation techniques) to ensure robust performance against technical variations further[18] (e-Method 4). Also, CTs of patients involved in the clinical evaluation groups could not be part of AI training to avoid overfitting.

Details on the previously published rationale and methodological choices to obtain ground truth labels, are available in e-Method 2. Briefly, the AI-driven algorithm allows for the holistic segmentation of five semantic labels, representative of the main structural hallmarks of CF: bronchiectasis, peribronchial thickening, bronchial mucus, bronchiolar mucus, and collapse/consolidation (Figure 2). The Total Abnormal Volume (TAV) corresponds to the sum of the five labels volume. In the NOVAA-CT system, TAV corresponds to the main outcome, while the component scores, derived from the five semantic labels, enable complementary characterizations of structural imaging profiles. A sixth label was also created, which corresponds to the surrounding lung parenchyma. The Total Lung Volume was defined as the sum of (Total Abnormal Volume + Lung Parenchyma Volume).

Definitions of structural abnormalities were those of the Fleischner Society[24], and bronchiectasis was defined by considering an inner airway-artery ratio $\geq$ 1[25, 26] (e-Method

2). To take into account variations in lung volumes, notably between children and adults, normalization was performed as follows:

Normalized Volume of Label = (Volume of the Label) / (Total Lung Volume) x 10<sup>4</sup> [15] (e-Method 2)

NOVAA-CT predictions were visually checked by two observers with 8 years (IB) and 12 years (GD) of experience. The predictions were included for further analysis only if they were considered “visually accurate”, as defined by a visual agreement of more than 80% that would be acknowledged, subjectively, by both observers, in agreement with the expected accuracy of the technique[15, 17].

Also, CT reconstructed with standard kernels were used to apply the AI algorithm, owing to their known robustness against noise and comparability across machine brands and manufacturers[27, 28].

#### *Visual analysis of the CT Bhalla score*

Two observers with 8 years (IB) and 12 years (GD) of experience in thoracic imaging analysed lung images reconstructed with hard kernels independently and the mean of evaluations was recorded. All examinations were anonymized and analysed in random order, by using the Bhalla scoring system[7]. Briefly, the scoring system describes the severity of peribronchial thickening, the extent and severity of bronchiectasis, the extent of mucus plugs, the bronchial generation involved, the extent of collapse/consolidation, bulla, sacculation/abscess, and emphysema, considering a total of 18 lung segments per patient (e-Table 2). The score ranges between 0 to 25, the latter indicating an absence of structural alterations.

#### *Evaluation of reproducibility*

All 219 CT scans in 139 CF patients were evaluated twice to assess the reproducibility of NOVAA-CT on two different computer machines (e-Method 5), and the post-processing time was recorded.

#### *Statistical analysis.*

Statistical analysis was carried out by using MedCalc® Software 20.218 Ltd. No assumption on the normality of data distribution was made. Quantitative data were expressed as median with [interquartile range] and categorical data as numbers and percentages. Comparisons of medians were done with the Mann-Whitney test and paired medians with the Wilcoxon test. For categorical variables, percentages were compared using the Chi-Square test. The relationship between NOVAA-CT, Bhalla scores, and PFT was assessed using Spearman correlation on ranks, per each independent group, and a p-value<0.05 was considered significant[29]. Reproducibility was assessed using

intraclass correlation coefficients (ICC) and Bland-Altman analysis[30]. Correlation values were scored as slight ( $\leq 0.20$ ), fair ( $\geq 0.20$  and  $< 0.40$ ), moderate ( $\geq 0.40$  and  $< 0.60$ ), good ( $\geq 0.60$  and  $< 0.80$ ), very good ( $\geq 0.80$  and  $< 0.95$ ) and almost perfect ( $\geq 0.95$ )[31, 32]. The minimum number of subjects per group was calculated to assess to good coefficient of correlation  $\geq 0.6$  between TAV and FEV1%p, with an alpha risk of 0.05 and a power of 0.20, indicating a minimum number of 19.

## RESULTS

### *Study populations*

A total of 139 CF patients were included across seven geographically-distinct Institutions, and composed of three independent groups (Table 1).

In the ETI group, 98 consecutive patients who had been treated with ETI were screened. These patients were followed up with imaging at 1 year by using either MRI or CT and 60 had completed both CT and PFT at 1-year intervals. The median age was 14.8 years, ranging from 12 to 50 years old.

For the ABPA group, patients were screened from the same Institution but from a temporally distinct period, and there was no patient overlap with the ETI group. Herein, 23 out of 137 CF patients were included and all of them were treated for ABPA using corticosteroids. However, three patients were excluded because a CT scan at 1 year was not available. The median age was 14 years, from 9- to 32-year-old and the median Bhalla score was 14.

Regarding the External group, 59 CF patients were included. The median age of patients in the external dataset was 25 years from 8- to 67-year-old and the median CT Bhalla score was 12.

All 139 NOVAA-CT predictions were considered accurate visually, and thus, included for analysis. No false positive depiction was noticed.

### *Baseline correlations between NOVAA-CT and established CF outcomes*

At cross-sectional analysis, there was a significant baseline correlation between TAV and FEV1%p ( $\rho=0.81$ ;  $\rho=0.81$ ,  $p<0.001$ ) and the Bhalla score ( $\rho=0.95$ ;  $\rho=0.78$ ;  $p<0.001$ ) in both ETI ( $n=60$ ) and ABPA ( $n=20$ ) groups, respectively.

When applying NOVAA-CT in a multicenter External group, these correlations were still found very good with FEV1%p ( $\rho=0.81$ ,  $p<0.001$ ) and good with the Bhalla score ( $\rho=0.77$ ,  $p<0.001$ ).

Also, significant correlations were observed regarding all NOVAA-CT sub-scores in all three groups ( $\rho\geq 0.60$ ;  $p\leq 0.005$ ) (Table 2).

### *Longitudinal changes of NOVAA-CT in CF patients undergoing ETI*

After treatment with a triple combination of Elexacaftor/Tezacaftor/Ivacaftor (n=60), there was a modification in disease status with an improvement in median FEV1%p, from 82.5 to 103 ( $p<0.001$ ) and median Bhalla score, from 14 to 16 ( $p<0.001$ ) (Table 3).

After such substantial improvement in disease severity status at M12, the correlations between TAV and FEV1% and Bhalla score remained good to very good ( $\rho=0.78$  and  $\rho=0.86$ , respectively;  $p<0.001$ ) (Table 2).

At paired-comparisons, median TAV decreased 67.3 to 7.6 ( $p<0.001$ ) and all sub-scores demonstrated a significant reduction in the volumes of structural abnormalities ( $p<0.001$ ), such as bronchiectasis, bronchial mucus and bronchiolar mucus (Table 3). Of note, the volumes of bronchiectasis decreased in 53 out of 60 CF patients with ETI treatment ( $p<0.001$ ) (Figures 3 and 4).

Moreover, FEV1% did not improve in 11 ETI patients, although normalized TAV decreased in all ETI patients.

Moreover, there was no significant change in the absolute value of the Total Lung Volume between M0 and M12 ( $p=0.15$ ; e-Table 3) or the mean lung density ( $p=0.09$ ; e-Table 1).

### *Longitudinal changes of NOVAA-CT in CF patients undergoing corticosteroids*

After treatment with corticosteroids for ABPA, there was an improvement in median FEV1%p from 82 to 83.5 ( $p=0.003$ ) and Bhalla score from 12 to 14 ( $p<0.001$ ) (Table 3). At M12, the correlations between NOVAA-CT and FEV1% and the CT Bhalla score remained significant ( $p\leq 0.005$ ) (Table 2).

This time, there was a significant reduction in median TAV from 146 to 34 ( $p<0.001$ ) (Table 3; e-Table 3) and almost all structural abnormalities volumes ( $p\leq 0.01$ ), except for bronchiectasis ( $p=0.14$ ) (Figure 5-6; e-Figure 1).

### *Reproducibility and acquisition time of NOVAA-CT*

The automated measurement were equal between repeat measurements ( $ICC>0.99$ ; mean difference = 0 [Limits of agreement: 0; 0]) in all groups and all time points. The mean processing time to get NOVAA-CT was 6 and 8 minutes *per* CT scan by using two different computers, respectively.

### *Correlation and reproducibility of the visual Bhalla score*

An additional analysis was performed to assess correlations and reproducibility of the visual Bhalla score. By considering the whole dataset of CT examinations, the correlation to FEV1%p was significant ( $\rho=0.70$ ;  $p<0.001$ ). The interobserver reproducibility between expert readers was very good (ICC=0.90; mean difference -0.5 [-2.1, 2.2]).

## **DISCUSSION**

The study demonstrates that NOVAA-CT can be a valid and responsive quantitative CT outcome in CF. Significant relationships to the disease status were consistently found in three independent clinical scenarios, including temporally-distinct longitudinal improvement with two different treatments, and a geographically-distinct multicenter group. Also, characterization of structural abnormality profiles was allowed for enabling differing characterization of treatment effects on potentially reversible structural abnormalities.

To our knowledge, reports on 3D holistic segmentation of airway lesions have been scarce yet, while evaluation of responsiveness with treatment have been lacking. Indeed, segmentations of such fine and complex structures are challenging to reach. Nevertheless, there is a growing need for developing tools to assess bronchiectasis[33] and CF-related airway lesions[3]. An initial feasibility study using a combination of 2D-CNNs to address this issue has been reported[15]; however, in this previous study, evaluation of responsiveness with lumacaftor/ivacaftor was tested in a small population of 10 CF patients only, and no multicenter evaluation was performed. More recently, the process has been moved towards a single 3D-nnUnet CNN architecture[17], demonstrating more accurate similarity performance as compared to its 2D counterpart. For instance, in 2D implementation, it is known that prediction inconsistencies can occur between contiguous CT slices because each 2D-CT slice is predicted independently. Conversely, 3D-CNN take into account their local environment[34]. Herein, we conceptualize the principle of using the 3D holistic segmentations to derive an automated quantitative outcome that we will call NOVAA-CT and we aimed at assessing both clinical validity and responsiveness.

Firstly, cross-sectional evaluation of NOVAA-CT validity showed strong correlations to two established clinical outcomes of CF, *i.e.* FEV1% and the CT Bhalla score, including temporally distinct and geographically distinct groups of evaluations[18]. Thus, the relationship to the CF disease severity status appeared consistent in all three independent groups, using different CT machines brands, over a wide range of ages, disease severities and a multicenter setting, supporting the generalizability of the technique.

Secondly, longitudinal evaluations of NOVAA-CT responsiveness demonstrated sensitive and reproducible detection of improvements in two different clinical conditions, using either CFTR modulators or corticosteroid treatments. In the field of ABPA, significant improvements were quantified for peribronchial thickening, bronchial mucus, and collapse/consolidation but not for bronchiectasis. Thus, these automated results appear consistent with those found in previous reports in non-CF ABPA condition[35], where a reduction in bronchial mucus but not bronchiectasis was observed. In addition, the lack of change of bronchial dilatation, in this group, appears in agreement with the classical view of bronchiectasis as an irreversible bronchial widening because of a definite scarring process[25].

In 60 CF patients undergoing ETI, it is noteworthy that TAV consistently improved in all patients at CT, although a more heterogenous response was found with FEV1%. In addition, the effects of treatment appeared different than in ABPA since all labels were found to decrease after treatment, including the bronchiectasis volume. This latter finding deserves further comments. Indeed, several reports have noticed, visually, the possibility of bronchiectasis reversibility with ETI [36–40]. In this study, NOVAA-CT quantified the degree of bronchial widening in 3D, with delineation of their complex contours at a voxel-wise level of precision. Conversely, visual scores may allow to diagnose the presence of bronchiectasis, but may lack ability to assess the degree of bronchial widening confidently, if not an inability[10]. By taking into account the holistic morphology in 3D, it is noteworthy that a quantitative reduction in bronchiectasis volume was found in the majority of these CF patients undergoing ETI. However, as compared to the literature in adults with advanced disease, where a reversibility of bronchiectasis was not noticed[41], our study group was mainly composed of young CF patients with mild-to-moderate disease severity[37]. Thus, this finding may advocate for early therapeutic intervention, before irreversible damages may occur. In addition, these results provide objective evidence that bronchiectasis reversibility is possible in the context of a primary airway disease, which is more likely to be observed recently with the advent of effective CFTR modulators. Next multicenter studies should be conducted to understand better whether reversal of bronchiectasis may be sustained by a distinct phenotype of CF patients.

Thirdly, as a quantitative CT outcome, NOVAA-CT differs from standard visual scoring methods in several ways. Indeed, visual methods often use categories to describe the CT information in broad terms, using a subset of CT slices, or scoring the disease severity by taking into account only the most predominant one[10, 42]. The range of variation is also narrow, especially regarding the component scores, making longitudinal evaluations often equivocal at the individual level[10]. Moreover, AI allows for human-like computer vision tasks, to reach semantic recognitions and understanding on which specific structural abnormalities a given treatment may have an effect. In addition, the

automated outcome was reproducible, without a need for trained expert readers but simple computer skills.

As compared to other quantitative methods, we could discuss that PRAGMA-CF has been showed sensitive to the disease variations[43]. However, this method is both time consuming and different by using a hierarchical scoring system. As a limitation, Yuchin Chen and colleagues recently reported that such system automatically under sample those abnormalities that do not have the highest hierarchical rank, in addition to the use of 10 CT slices only[16]. Other methods such as airway-artery ratio or standard measurements of bronchial surfaces have been reported[16, 44], as the mean of several 1D or 2D measurements. However, for technical reasons, these softwares are excluding all bronchial paths distally from airway occlusion by mucus plugging. Hence, Dettmer and colleagues reported that longitudinal evaluations are equivocal when different bronchial datasets are used to calculate the mean values over time[44]. Conversely, NOVAA-CT provides a non-hierarchical but comprehensive evaluation, over the entire lung. NOVAA-CT integrate all voxels to quantify the actual 3D extent of all airway lesions, without averaging, exclusion of bronchial path or selection of abnormalities due to a hierarchical ranking. Nevertheless, comparison of the NOVAA-CT holistic evaluation to other recent methods was out of the scope of this validation study.

Next studies should be worth doing to explore the ability of NOVAA-CT in phenotyping CF patients according to differing label profiles. Translation to other non-CF conditions such as asthma or COPD should be worth evaluating. Moreover, there is a growing research interest regarding bronchiectasis[33, 45], in both CF and non-CF conditions, for which NOVAA-CT could offer novel tools of investigation to speed up research discoveries and understanding of the disease mechanisms.

The study had several limitations. The design was retrospective, and prospective evaluations may be worth doing[18]. Treatments effects on structural imaging were evaluated in CF patients from a single Institution, and the monocenter findings should be evaluated in a multicenter setting further, at best during an international study. The minimum age was 8 years, and evaluation in preschool children was not performed. The multilabeling process was not trained to segment air trapping or lung vessels[46]. A single cut-off to define bronchiectasis was implemented, and next implementation could allow the user to test alternative cut-offs. Finally, lobar or segmental lung segmentations were not implemented for a deeper insight into the regional distributions of the disease.

To conclude, NOVAA-CT automated scoring demonstrates validity, reliability and responsiveness for monitoring CF disease severity over an entire lung, and quantifies therapeutic effects on lung structure at CT, such as the volumetric reversibility of airway abnormalities with ETI.

## References

1. Cromwell EA, Ostrenga JS, Todd JV, et al (2023) Cystic fibrosis prevalence in the United States and participation in the Cystic Fibrosis Foundation Patient Registry in 2020. *J Cyst Fibros* 22:436–442. <https://doi.org/10.1016/j.jcf.2023.02.009>
2. Middleton PG, Mall MA, Dřevínek P, et al (2019) Elexacaftor-Tezacaftor-Ivacaftor for Cystic Fibrosis with a Single Phe508del Allele. *N Engl J Med* 381:1809–1819. <https://doi.org/10.1056/NEJMoa1908639>
3. Heltshe SL, Cogen J, Ramos KJ, Goss CH (2017) Cystic Fibrosis: The Dawn of a New Therapeutic Era. *Am J Respir Crit Care Med* 195:979–984. <https://doi.org/10.1164/rccm.201606-1250PP>
4. Tiddens HAWM, Kuo W, van Straten M, Ciet P (2018) Paediatric lung imaging: the times they are a-changin'. *Eur Respir Rev* 27:170097. <https://doi.org/10.1183/16000617.0097-2017>
5. Ramsey BW (2007) Use of lung imaging studies as outcome measures for development of new therapies in cystic fibrosis. *Proc Am Thorac Soc* 4:359–363. <https://doi.org/10.1513/pats.200611-183HT>
6. Zorzo C, Caballero P, Diab L, et al (2020) Predictive value of computed tomography scoring systems evolution in adults with cystic fibrosis. *Eur Radiol* 30:3634–3640. <https://doi.org/10.1007/s00330-020-06759-z>
7. Bhalla M, Turcios N, Aponte V, et al (1991) Cystic fibrosis: scoring system with thin-section CT. *Radiology* 179:783–788. <https://doi.org/10.1148/radiology.179.3.2027992>
8. Brody AS, Kosorok MR, Li Z, et al (2006) Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. *J Thorac Imaging* 21:14–21. <https://doi.org/10.1097/01.rti.0000203937.82276.ce>
9. Diab-Cáceres L, Girón-Moreno RM, García-Castillo E, et al (2021) Predictive value of the modified Bhalla score for assessment of pulmonary exacerbations in adults with cystic fibrosis. *Eur Radiol* 31:112–120. <https://doi.org/10.1007/s00330-020-07095-y>
10. Calder AD, Bush A, Brody AS, Owens CM (2014) Scoring of chest CT in children with cystic fibrosis: state of the art. *Pediatr Radiol* 44:1496–1506. <https://doi.org/10.1007/s00247-013-2867-y>
11. Benlala I, Point S, Leung C, et al (2020) Volumetric quantification of lung MR signal intensities using ultrashort TE as an automated score in cystic fibrosis. *Eur Radiol* 30:5479–5488. <https://doi.org/10.1007/s00330-020-06910-w>
12. Chassagnon G, Martin C, Burgel P-R, et al (2018) An automated computed tomography score for the cystic fibrosis lung. *Eur Radiol* 28:5111–5120. <https://doi.org/10.1007/s00330-018-5516-x>
13. DeBoer EM, Swiercz W, Heltshe SL, et al (2014) Automated CT scan scores of bronchiectasis and air trapping in cystic fibrosis. *Chest* 145:593–603. <https://doi.org/10.1378/chest.13-0588>
14. Rosenow T (2018) Quantification of CT bronchiectasis and its relationship to ventilation in cystic fibrosis. *Thorax* 73:4. <https://doi.org/10.1136/thoraxjnl-2017-210917>

15. Dournes G, Hall CS, Willmering MM, et al (2022) Artificial intelligence in computed tomography for quantifying lung changes in the era of CFTR modulators. *Eur Respir J* 59:2100844. <https://doi.org/10.1183/13993003.00844-2021>
16. Chen Y, Lv Q, Andrinopoulou E-R, et al (2023) Automatic bronchus and artery analysis on chest computed tomography to evaluate the effect of inhaled hypertonic saline in children aged 3-6 years with cystic fibrosis in a randomized clinical trial. *J Cyst Fibros* 22:916–925. <https://doi.org/10.1016/j.jcf.2023.05.013>
17. Hadj Bouzid AI, Denis de Senneville B, Baldacci F, et al (2024) CT evaluation of 2D and 3D holistic deep learning methods for the volumetric segmentation of airway lesions. 2024 IEEE International Symposium on Biomedical Imaging (ISBI). <https://doi.org/10.48550/ARXIV.2403.08042>
18. Park SH, Han K (2018) Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology* 286:800–809. <https://doi.org/10.1148/radiol.2017171920>
19. Mokkink L, Terwee C, De Vet H (2021) Key concepts in clinical epidemiology: Responsiveness, the longitudinal aspect of validity. *Journal of Clinical Epidemiology* 140:159–162. <https://doi.org/10.1016/j.jclinepi.2021.06.002>
20. PND5 mucoviscidose [Internet]. Available from: [https://www.has-sante.fr/upload/docs/application/pdf/2017-09/pnds\\_2017\\_vf1.pdf](https://www.has-sante.fr/upload/docs/application/pdf/2017-09/pnds_2017_vf1.pdf). Last accessed March 1, 2024.
21. Stevens DA, Moss RB, Kurup VP, et al (2003) Allergic bronchopulmonary aspergillosis in cystic fibrosis--state of the art: Cystic Fibrosis Foundation Consensus Conference. *Clin Infect Dis* 37 Suppl 3:S225-264. <https://doi.org/10.1086/376525>
22. Miller MR, Hankinson J, Brusasco V, et al (2005) Standardisation of spirometry. *Eur Respir J* 26:319–338. <https://doi.org/10.1183/09031936.05.00034805>
23. Wanger J, Clausen JL, Coates A, et al (2005) Standardisation of the measurement of lung volumes. *Eur Respir J* 26:511–522. <https://doi.org/10.1183/09031936.05.00035005>
24. Bankier AA, MacMahon H, Colby T, et al (2024) Fleischner Society: Glossary of Terms for Thoracic Imaging. *Radiology* 310:e232558. <https://doi.org/10.1148/radiol.232558>
25. Aliberti S, Goeminne PC, O'Donnell AE, et al (2022) Criteria and definitions for the radiological and clinical diagnosis of bronchiectasis in adults for use in clinical trials: international consensus recommendations. *Lancet Respir Med* 10:298–306. [https://doi.org/10.1016/S2213-2600\(21\)00277-0](https://doi.org/10.1016/S2213-2600(21)00277-0)
26. Meerburg JJ, Veerman GDM, Aliberti S, Tiddens HAWM (2020) Diagnosis and quantification of bronchiectasis using computed tomography or magnetic resonance imaging: A systematic review. *Respir Med* 170:105954. <https://doi.org/10.1016/j.rmed.2020.105954>
27. Mackin D, Ger R, Gay S, et al (2019) Matching and Homogenizing Convolution Kernels for Quantitative Studies in Computed Tomography. *Invest Radiol* 54:288–295. <https://doi.org/10.1097/RLI.0000000000000540>
28. Solomon JB, Christianson O, Samei E (2012) Quantitative comparison of noise texture across CT scanners from different manufacturers. *Med Phys* 39:6048–6055. <https://doi.org/10.1118/1.4752209>

29. Streiner DL, Norman GR (2011) Correction for multiple testing: is there a resolution? *Chest* 140:16–18. <https://doi.org/10.1378/chest.11-0523>
30. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
31. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
32. Fleiss JL, Cohen J (1973) The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement* 33:613–619. <https://doi.org/10.1177/001316447303300309>
33. Mazulov O, Hill AT, Marchant J (2023) Developments and priorities in bronchiectasis research. *The Lancet Respiratory Medicine* 11:669–670. [https://doi.org/10.1016/S2213-2600\(23\)00258-8](https://doi.org/10.1016/S2213-2600(23)00258-8)
34. Niyas S, Pawan SJ, Anand Kumar M, Rajan J (2022) Medical image segmentation with 3D convolutional neural networks: A survey. *Neurocomputing* 493:397–413. <https://doi.org/10.1016/j.neucom.2022.04.065>
35. Godet C, Brun A-L, Couturaud F, et al (2024) CT Imaging Assessment of Response to Treatment in Allergic Bronchopulmonary Aspergillosis in Adults With Bronchial Asthma. *Chest* S0012-3692(24)00261–7. <https://doi.org/10.1016/j.chest.2024.02.026>
36. Gushue C, Eisner M, Bai S, et al (2023) Impact of Elexacaftor–Tezacaftor–Ivacaftor on lung disease in cystic fibrosis. *Pediatric Pulmonology* ppul.26485. <https://doi.org/10.1002/ppul.26485>
37. David M, Benlala I, Bui S, et al (2023) Longitudinal Evaluation of Bronchial Changes in Cystic Fibrosis Patients Undergoing Elexacaftor/Tezacaftor/Ivacaftor Therapy Using Lung MRI With Ultrashort Echo-Times. *Magnetic Resonance Imaging* jmri.29041. <https://doi.org/10.1002/jmri.29041>
38. Middleton PG, Simmonds NJ (2023) Cystic fibrosis modulator therapy can reverse cystic bronchiectasis. *Respirol Case Rep* 11:e01172. <https://doi.org/10.1002/rcr2.1172>
39. Chang AB, Kantar A, Redding GJ, Hill AT (2022) Is Bronchiectasis a Reversible Disorder? *Arch Bronconeumol* S0300-2896(22)00633–0. <https://doi.org/10.1016/j.arbres.2022.11.007>
40. Madrid Carbajal C, Palomo B, García Clemente M (2023) Radiological Changes After Treatment With Triple Therapy in Cystic Fibrosis. Are Bronchiectasis Irreversible? *Arch Bronconeumol* 59:118. <https://doi.org/10.1016/j.arbres.2022.10.011>
41. Bec R, Reynaud-Gaubert M, Arnaud F, et al (2022) Chest computed tomography improvement in patients with cystic fibrosis treated with elexacaftor-tezacaftor-ivacaftor: Early report. *Eur J Radiol* 154:110421. <https://doi.org/10.1016/j.ejrad.2022.110421>
42. van Straten M, Brody AS, Ernst C, et al (2020) Guidance for computed tomography (CT) imaging of the lungs for patients with cystic fibrosis (CF) in research studies. *J Cyst Fibros* 19:176–183. <https://doi.org/10.1016/j.jcf.2019.09.001>
43. Rosenow T, Oudraad MCJ, Murray CP, et al (2015) PRAGMA-CF. A Quantitative Structural Lung Disease Computed Tomography Outcome in Young Children with Cystic Fibrosis. *Am J Respir Crit Care Med* 191:1158–1165. <https://doi.org/10.1164/rccm.201501-0061OC>
44. Dettmer S, Weinheimer O, Sauer-Heilborn A, et al (2023) Qualitative and quantitative evaluation of computed tomography changes in adults with cystic fibrosis treated with elexacaftor-

tezacaftor-ivacaftor: a retrospective observational study. *Front Pharmacol* 14:1245885. <https://doi.org/10.3389/fphar.2023.1245885>

45. Chalmers JD, Aksamit T, Aliberti S, et al (2022) World Bronchiectasis Day 2022. *Eur Respir J* 59:2201249. <https://doi.org/10.1183/13993003.01249-2022>
46. Nam JG, Witanto JN, Park SJ, et al (2021) Automatic pulmonary vessel segmentation on noncontrast chest CT: deep learning algorithm developed using spatiotemporally matched virtual noncontrast images and low-keV contrast-enhanced vessel maps. *Eur Radiol* 31:9012–9021. <https://doi.org/10.1007/s00330-021-08036-z>

## TABLES

**Table 1. Characteristics of CF patients in three independent groups.**

		ETI group (n=60)		ABPA group (n=20)		External group (n=59)	
<b>Age</b>	Years	14.8	[13-24]	14	[11-18]	25	[14-33]
<b>Sex</b>	Male / Female	32 / 28		11 / 9		28 / 31	
<b>BMI</b>	(kg.m <sup>-2</sup> )	21	[19-23]	20	[16-21]	20	[17-22]
<b>DeltaF508 variation</b>	Homozygous	42		12		33	
	Heterozygous	28		8		23	
	Other	0		0		3	
<b>Immunology</b>	Total IgE (UI/ml)	80	[25-150]	940	[430-2000]	-	
<b>Pulmonary Function Test</b>	FEV1%p	82.5	[59-102]	82	[51-94]	72	[48-92]
	FEV1/FVC	74	[60-86]	78	[68-86]	83	[68-95]
	RV%p	107	[87-154]	110	[80-178]	121	[88-151]
	TLC%p	100	[86-108]	95	[84-104]	100	[84-106]
<b>Computed Tomography</b>	Bhalla score	14	[12-16]	12	[10-14]	12	[11-15]

Footnote: Data are medians with [interquartile range] for continuous data and absolute numbers for categorical data.

Legends: CF=cystic fibrosis; ETI=Elexacaftor/Tezacaftor/Ivacaftor; ABPA=allergic broncho-pulmonary aspergillosis; IgE=immunoglobulin E; BMI=body mass index; FEV1=forced expiratory volume in 1-second percentage predicted; FVC=forced vital capacity; RV=residual volume; TLC=total lung capacity; %p=percentage predicted.

**Table 2. Correlations between quantitative CT measurements, pulmonary function test and the visual Bhalla score.**

NOVAA-CT		ETI group (n=60)				ABPA group (n=20)				External group (n=59)	
		Bhalla Score		FEV1%p		Bhalla Score		FEV1%p		Bhalla Score	FEV1%p
		M0	M12	M0	M12	M0	M12	M0	M12	M0	M0
<b>Bronchiectasis</b>	$\rho$	-0.89	-0.82	-0.83	-0.70	-0.77	-0.81	-0.88	-0.89	-0.74	-0.66
	p-value	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>
<b>Peribronchial thickening</b>	$\rho$	-0.92	-0.85	-0.78	-0.78	-0.86	-0.82	-0.80	-0.80	-0.75	-0.74
	p-value	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>
<b>Bronchial mucus</b>	$\rho$	-0.84	-0.77	-0.75	-0.75	-0.71	-0.87	-0.63	-0.82	-0.78	-0.67
	p-value	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>0.002</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>
<b>Bronchiolar mucus</b>	$\rho$	-0.82	-0.77	-0.77	-0.78	-0.69	-0.80	-0.70	-0.86	-0.70	-0.62
	p-value	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>
<b>Collapse/Consolidation</b>	$\rho$	-0.76	-0.68	-0.67	-0.69	-0.63	-0.61	-0.60	-0.63	-0.64	-0.64
	p-value	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>0.002</b>	<b>0.004</b>	<b>0.005</b>	<b>0.002</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>
<b>Total Abnormal Volume</b>	$\rho$	-0.95	-0.86	-0.81	-0.76	-0.78	-0.82	-0.81	-0.77	-0.81	-0.77
	p-value	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>0.005</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>

Footnote: Correlations were assessed per each independent ETI, ABPA and External groups and timepoints with rho coefficients of Spearman on ranks. A p-value<0.05 was considered significant as indicated in bold characters.

Legends: NOVAA=normalized volume of airway abnormalities; CT=computed tomography; M0=initial evaluation; M12=second evaluation at 1 year after treatment; ETI=Elexacaftor/Tezacaftor/Ivacaftor; ABPA=allergic broncho-pulmonary aspergillosis; FEV1%p=forced expiratory volume in 1-second percentage predicted.

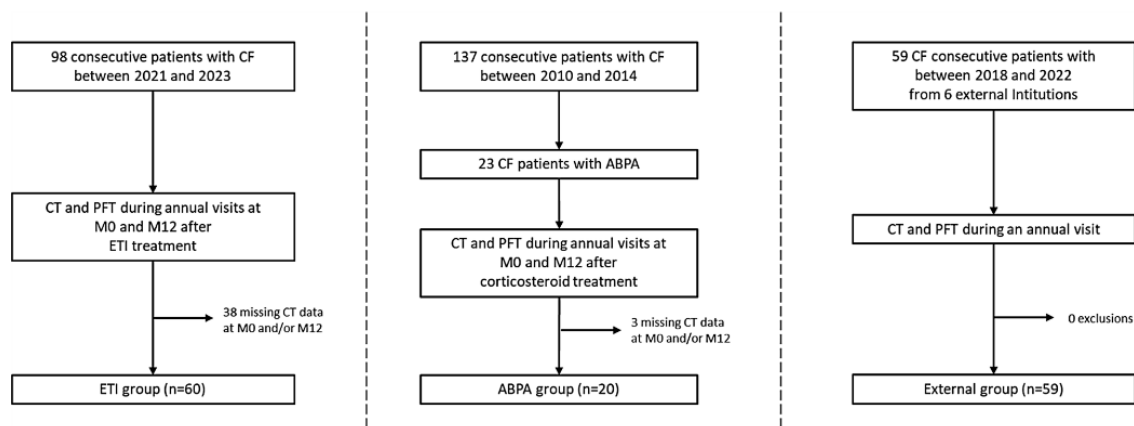
**Table 3. Longitudinal comparisons of quantitative CT measurements in CF patients undergoing either ETI or corticosteroid treatment.**

	ETI group (n = 60)					ABPA group (n = 20)				
	M0		M12		p-value	M0		M12		p-value
	Median	IQR	Median	IQR		Median	IQR	Median	IQR	
<b>Pulmonary function test</b>										
FEV1%p	82.5	[59-102]	103	[71-118]	<b>&lt;0.001</b>	82	[51-94]	83.5	[59-107]	<b>0.003</b>
<b>CT Bhalla score</b>										
Total Bhalla score	14	[12-16]	16	[13-18]	<b>&lt;0.001</b>	12	[10-14]	14	[13-17]	<b>&lt;0.001</b>
<b>NOVAA-CT</b>										
Bronchiectasis	15.8	[5.6-30]	3.4	[0.5-27]	<b>&lt;0.001</b>	13	[4-42]	15	[1-55]	0.14
Peribronchial thickening	30.3	[11-81]	3.5	[1-13]	<b>&lt;0.001</b>	42	[8-94]	6.5	[1-33]	<b>0.01</b>
Bronchial mucus	1.9	[0-43]	0	[0-0]	<b>&lt;0.001</b>	18	[1-65]	0.1	[0-2.7]	<b>&lt;0.001</b>
Bronchiolar mucus	5.9	[0.2-27]	0	[0-0]	<b>&lt;0.001</b>	7.8	[0.2-27]	1.1	[0-4.7]	<b>&lt;0.001</b>
Collapse/Consolidation	0.1	[0-1.3]	0	[0-0]	<b>&lt;0.001</b>	1	[0.0-36]	0.2	[0.0-4.3]	<b>0.009</b>
Total abnormal volume	67.3	[21-211]	7.6	[2-47]	<b>&lt;0.001</b>	146	[53-342]	34	[5.4-106]	<b>&lt;0.001</b>

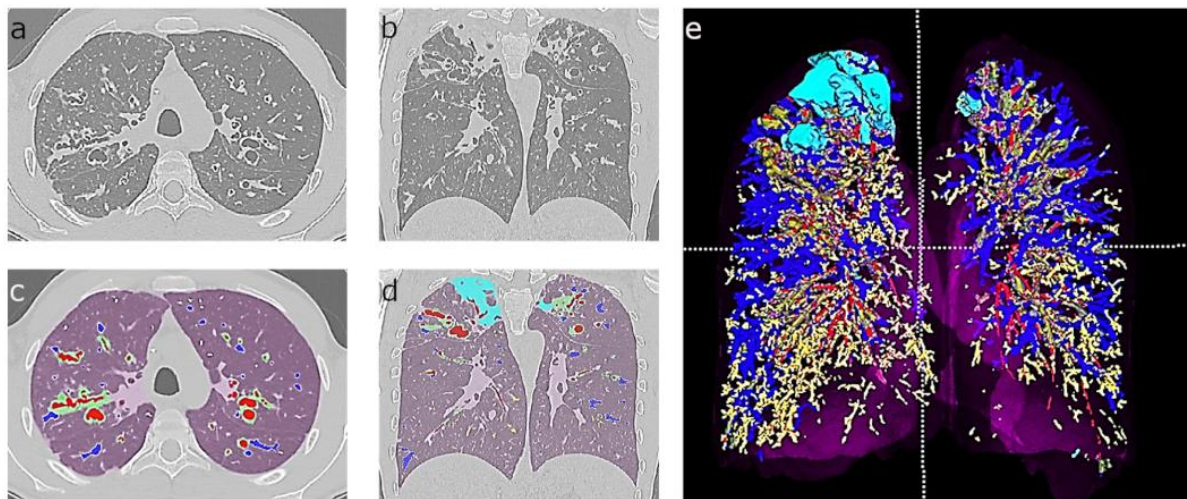
Footnote: Data are medians with [IQR: interquartile range]. A p-value<0.05 was considered significant as indicated in bold characters.

Legends: NOVAA=normalized volume of airway abnormalities; CT=computed tomography; M0=initial evaluation; M12=second evaluation at 1 year after treatment; ETI=Eleacaftor/Tezacaftor/Ivacaftor; ABPA=allergic broncho-pulmonary aspergillosis; FEV1%p=forced expiratory volume in 1-second percentage predicted.

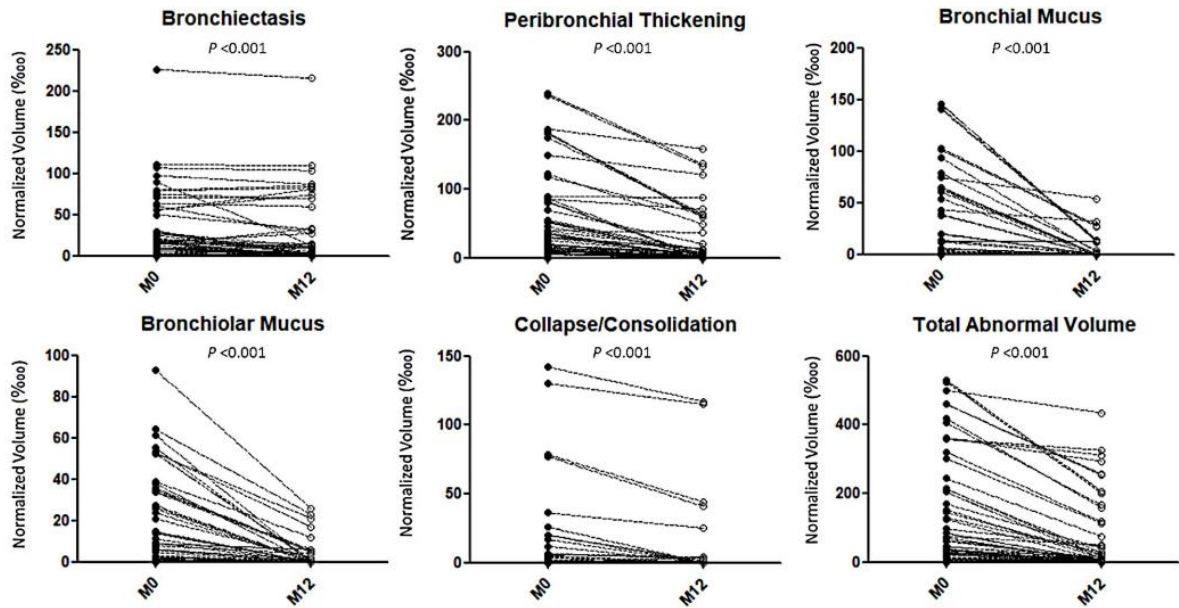
## FIGURES



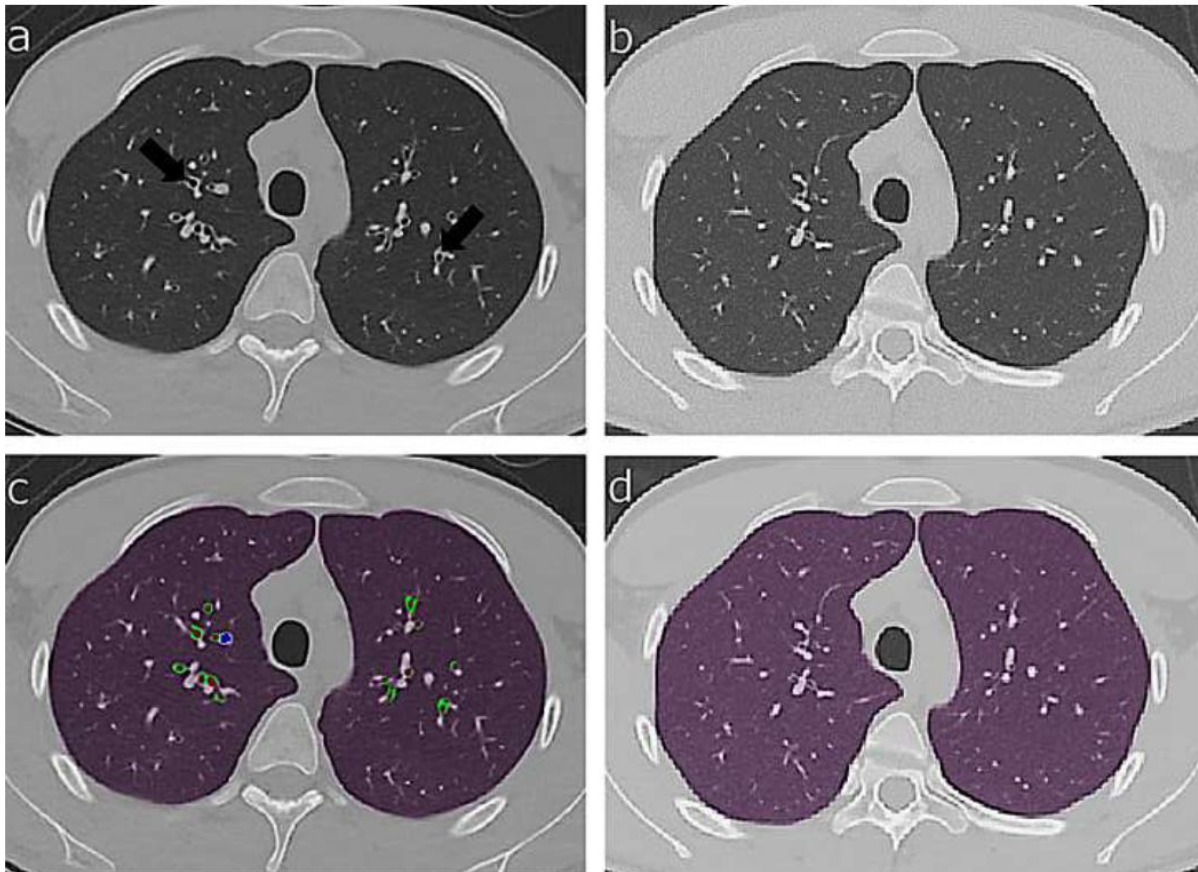
**Figure 1. Study flow chart.** CF=cystic fibrosis; CT=computed tomography; PFT=pulmonary function test; ETI=Elexacaftor/Tezacaftor/Ivacaftor; ABPA=allergic broncho-pulmonary aspergillosis; M0=initial evaluation; M12=second evaluation at 1 year.



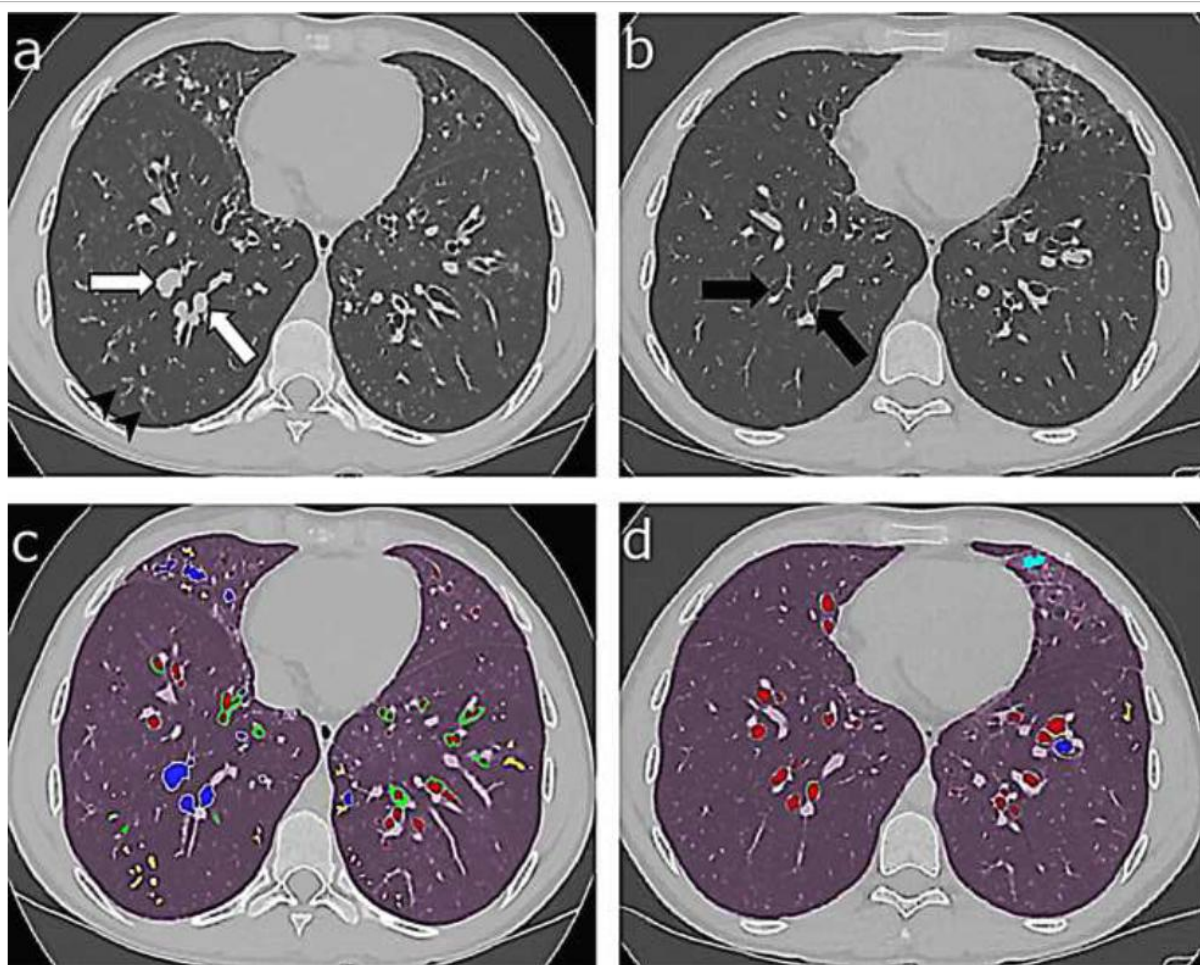
**Figure 2. Example of NOVA-CT quantitative measurements.** CT of the chest in a 21-year-old male with cystic fibrosis, in axial (a) and coronal reformation (b). On corresponding (c) and (d) images, holistic three-dimensional segmentation by deep learning is displayed, with automated labeling of bronchiectasis (red color), peribronchial thickening (green color), bronchial mucus (blue color), bronchiolar mucus (yellow color), collapse/consolidation (cyan color), and the lung parenchyma (purple color). 3D volume rendering of lesions is displayed in (e), showing the ability to calculate the volume of airway lesions over the full CT dataset. This patient had a severe disease status, with a FEV1% of 56% and a total abnormal volume of 508 ml.



**Figure 3. Longitudinal changes of NOVA-CT measurements in cystic fibrosis patients undergoing Elexacaftor/Tezacaftor/Ivacaftor treatment.** Black circles indicate normalized volume measurements at the initiation of treatment (M0) and white circles the corresponding values at one year (M12). A significant reduction in all label volumes was noticed at paired analysis between M0 and M12, including a reduction in bronchiectasis.

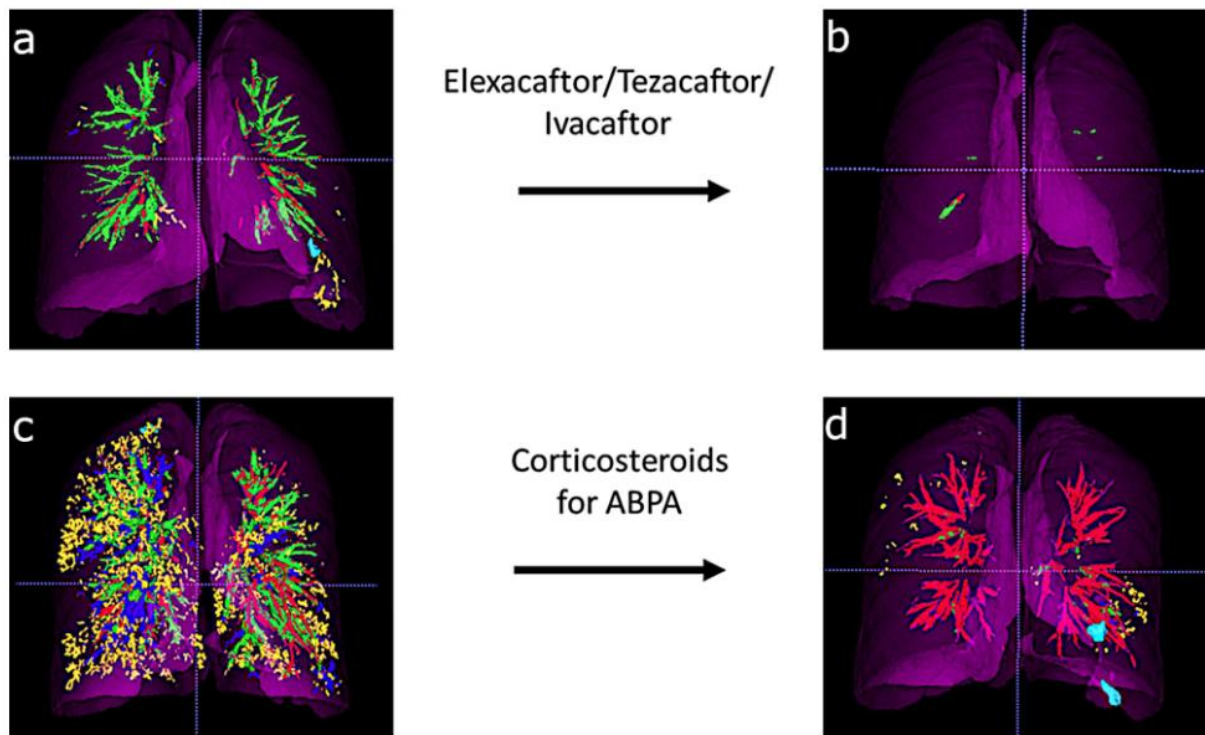


**Figure 4. Axial CT slice of the chest in a 14-year-old female with cystic fibrosis and undergoing Elexacaftor/Tezacaftor/Ivacaftor treatment at initiation (a) and after one year (b).** Low dose CT (1.2 mSv) was performed on a GE Revolution machine, using 120 kV, 150 mAs and reconstructed with a standard kernel. Black arrows indicate dilatation of the bronchial lumen (a), which is reversible after treatment (b). In corresponding (c) and (d) images, NOVAA-CT fully automated and holistic segmentations are shown. Red, green, blue, and purple colors indicate bronchiectasis, peribronchial thickening, bronchial mucus, and lung parenchyma, respectively. Note the complete disappearance of all labels on (d) image, indicating the deep learning algorithm predicted a complete resolution of all structural abnormalities after treatment.



**Figure 5. Axial CT slice of the chest in a 16-year-old male with cystic fibrosis and undergoing corticosteroid for allergic broncho-pulmonary aspergillosis treatment at initiation (a) and after one year (b).** Low dose CT (1.4 mSv) was done on a Siemens Definition64 machine brand, using 110 kV, 200 mAs, and reconstructed with a B40s kernel. White arrows indicate bronchoceles (a), black arrowhead small bronchiolar impaction and black arrows indicate mucus-free dilatation of the bronchial lumen (b). In corresponding (c) and (d) images, NOVAA-CT fully automated and holistic segmentations are shown. Red, green, blue, yellow, and purple colors indicate bronchiectasis,

peribronchial thickening, bronchial mucus, bronchiolar mucus, and lung parenchyma, respectively. Note the resolution of peribronchial thickening, bronchial, and bronchiolar mucus but not bronchiectasis after treatment.



**Figure 6. Examples of NOVA-CT holistic segmentations of displayed in 3D, demonstrating volumetric modifications of cystic fibrosis-related airway lesions in a 14-year-old female (a, b) and a 16-year-old male (c, d). Volumes in red, green, blue, yellow, cyan, and purple colors indicate bronchiectasis, peribronchial thickening, bronchial mucus, bronchiolar, mucus, collapse/consolidation, and the lung parenchyma, respectively. Using Elexacaftor/Tezacaftor/Ivacaftor (ETI) (a, b), there was an improvement of FEV1% from 81 to 102%, and using corticosteroids for allergic bronchopulmonary aspergillosis (ABPA) (c, d), there was an improvement in FEV1% from the 89% to 92%. Note the neat reduction in all labels' volumes with ETI (a, b), including a reversibility in bronchiectasis (red volume). Conversely, using corticosteroids for ABPA (c, d), there was a reduction in mucus but a lack of change in bronchiectasis volume.**