UNIVERSITÀ DEGLI STUDI DI MILANO

PhD course in *Epidemiology, Environment and Public Health*

XXXIII cycle

Department of Clinical Sciences and Community Health

Laboratory of Medical Statistics and Biometry "Giulio A. Maccacaro"

THE ASSOCIATION OF CLINICAL PROGNOSTIC FACTORS WITH VENTILATION INHOMOGENEITY: SINGLE CENTRE ANALYSIS OF PATIENTS WITH CYSTIC FIBROSIS FROM 2014 TO 2019.

Simone GAMBAZZA

MATRICOLA R12064

Supervisor Prof. Federico AMBROGI

PhD Course Coordinator: Prof. Carlo LA VECCHIA

A.A. 2019/2020

Preface

This PhD project was undertaken in collaboration with the Cystic Fibrosis Regional Reference Centre of Milan (Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico), directed by Professor Carla Colombo.

The main focus of this exploratory analysis is to evaluate the relationship of one marker of lung disease in Cystic Fibrosis (CF), namely lung clearance index ($LCI_{2.5}$), with some of the variables currently used in clinic use to assess CF disease, including well known prognostic factors. Understanding if $LCI_{2.5}$ is associated or not to these routinely measures used in CF centres worldwide, would allow the CF team to better evaluate the disease stage, especially in young patients, who cannot perform spirometry or show normal/above-normal forced expiratory volume in the first second (FEV_1), the gold standard outcome used to assess severity of lung disease in older CF patients. The second part of this analysis evaluates the possibility to generate phenotypes based on the same clinical characteristics and prognostic factors collected for the current project, and to evaluate how CF lung disease evolves in different phenotypes in terms of $LCI_{2.5}$ variation. Eventually, association of $LCI_{2.5}$ with time to pulmonary exacerbation is investigated as well.

The emerging role of $LCI_{2.5}$ in the assessment of CF lung disease and its potential as an earlier marker forced the CF team to understand its behaviour under multiple clinical perspectives. Within this project, new information about $LCI_{2.5}$ is collected, contributing to the decision whether including $LCI_{2.5}$ as an outcome measure into regular follow-up at the CF centre of Milano.

Abstract

BACKGROUND. Especially in children, spirometry may not capture the underlying lung disease, which has been described in small airways since young ages in cystic fibrosis (CF). Lung Clearance Index ($LCI_{2.5}$) is one of the derivatives obtained by the multiple breath washout ($MBWN_2$) test, recognized as a sensitive alternative for the assessment of lung function. The potential of this outcome in clinical setting has not been fully explored yet.

METHODS. Data were collected for each patient's visit from 2014 up to September 2019 from the CF Centre of Milano and analyzed according to different perspectives. The association between lung function measures and selected prognostic variables was investigated by means of ordinary least-square regression models; disease phenotypes were obtained according to children's clinical characteristics by means of agglomerative nesting hierarchical clustering; longitudinal association between lung function and selected variables, and phenotypes, was assayed fitting linear-mixed effect models; recurrence of pulmonary exacerbations (PE) was evaluated by extensions of the Cox proportional hazard models.

RESULTS. A total of 433 $MBWN_2$ measurements from 245 children with CF were analyzed. Cross-sectionally, both $LCI_{2.5}$ and FEV_1 %predicted showed significant association with CF clinical prognostic variables, i.e., *Pseudomonas aeruginosa* infection, pancreatic function, BMI Z-score and age, whereas only $LCI_{2.5}$ was associated with genotype.

Two different phenotypes of children around 11 years old emerged, differentiating one from another also by a mean $LCI_{2.5}$ difference of 3.37 (95%CI: 2.57 to 4.16) and by a mean FEV_1 difference of 1.2 (95%CI:0.7 to 1.6) Z-score. Children with the best phenotype were mainly characterized by absence of *F508del* mutation and by a normal pancreatic function.

Longitudinally, $LCI_{2.5}$ and FEV_1 % predicted remained stable over time and showed meaningful differences between phenotypes at each follow-up visit. However, infection by *Pseudomonas aeruginosa* was associated with longitudinal variation of $LCI_{2.5}$ only, showing that being free from *Pseudomonas aeruginosa* lowers $LCI_{2.5}$ by 0.82 (95%CI: -1.36 to -0.27).

Eventually, mean number of PE in the present cohort was 1.8, varying from 0 to 4. Children with higher $LCI_{2.5}$ are expected to experience a 6% higher risk of PE recurrence during their follow-up (HR 1.06, 95%CI 1.01 to 1.10), whereas FEV_1 % predicted did not show any evidence of association. An association with PE recurrence was also detected between phenotypes in favour of children with better clinical status (HR 0.46, 95%CI 0.34 to 0.60).

CONCLUSIONS

$LCI_{2.5}$ shows interesting associations with clinical characteristics not shared with FEV_1, %predicted, both in cross-sectional and longitudinal analyses. Clustering has shown a disease profile of children who share negative clinical prognostic factors, also in terms of ventilation inhomogeneity, and are at higher risk of recurrent pulmonary exacerbations. Altogether these findings add to the available literature in confirming the clinical utility of $MBWN_2$.

Background

Cystic fibrosis is a classic Mendelian autosomal recessive disorder that mostly affects the respiratory and gastrointestinal system. It is the most common genetic life shortening disorder in the Caucasian population and it is caused by a mutation in the gene encoding for Cystic Fibrosis Transmembrane conductance Regulator (CFTR). The CFTR protein transports chloride ions across the membrane of epithelial cells and regulates the transport of other ions. [1] CF has been known for centuries: European literature from the XVIII century warns that the child whose forehead tastes salty will die soon. However, the full clinical spectrum of the disease was only described in 1938 by Dorothy H. Andersen, who described cystic fibrosis of the pancreas in 49 patients and the disorder was subsequently associated with lung infection and salt loss during a heat wave in New York. [2] By 1958, Paul di Sant'Agnese discovered abnormalities in sweat electrolytes and a sweat test was designed soon after. Lap-Chee Tsui and colleagues identified the gene responsible for the disease in 1989; designing a gene therapy followed as a consequence.

The U.S. Cystic Fibrosis Foundation's projected life expectancy for patients has increased from 31 years to 37 years over the past decade, and a UK model predicting that a child born with cystic fibrosis today will typically live to be 50 years of age or more seems to be realistic.[2]

Epidemiology

Cystic fibrosis has a prevalence of approximately 1 in 3000 live births in the northern European descent, but the number decreases in African Americans and reaches very low values in the Asian population. In Western countries cystic fibrosis is usually diagnosed within a few weeks from birth in the context of newborn screening; despite this, 10% of patients, often those with a milder disease are still identified as adults. [3] The absence of phenylalanine at position 508, known as *F508del*,

which is class II mutation, accounts for about two-thirds of mutated alleles in northern European and North American populations. Patients who carry two *F508del* mutations are often described as having classic CF, but other combinations of mutations may cause the same degree of disease. In Europe the allele frequency still varies from 24.27% in Israel to 82.24% in Denmark.[4]

Aetiology

Cystic fibrosis is caused by mutations in the *CFTR* gene (located on chromosome 7), which produces a glycoprotein, part of the ATP-binding cassette superfamily. Over 2000 variants in the *CFTR* gene have been discovered and grouped into six classes according to the functional defect. [5] Class I (i.e., *G542X*), II (i.e., *F508del*), and III mutations (i.e., *G551D*) are associated with no residual *CFTR* function and patients with these mutations have a severe phenotype. Class II mutations such as *F508del* lead to misfolding, premature degradation by the endoplasmic reticulum quality-control system, and impaired protein biogenesis, severely reducing the number of *CFTR* molecules that reach the cell surface.[6] By contrary, individuals with class IV (i.e., *R117H*), V (i.e., *3849+10kbC→T*) and VI (i.e., *4326delTC*) mutations conserve some residual function on *CFTR* protein. On average, they show a mild lung phenotype and pancreatic sufficiency, [1, 7] since *CFTR* variants are actually expressed at the apical membrane of secretory epithelia. Generally, some mutations seem to be of little clinical relevance yet.

The *CFTR* protein is a cAMP-regulated ion channel expressed on epithelial cells, composed of 12 alpha helices spanning the cell membrane and two ATP hydrolysing domains. Its main activity is transporting chloride ions, but it has also other regulatory functions, such as down-regulation of the amiloride-sensitive epithelial Na+ channel (ENaC) and bicarbonate-chloride exchange. [1]

Polymorphisms in modifiers genes (i.e. growth factor β1 and mannose-binding lectin 2) could explain the variability in the presentation and severity of the disease, as could environmental and socioeconomic factors. [1]

Pathophysiology

There are several hypotheses regarding how *CFTR* dysfunction leads to the phenotypic disease known as cystic fibrosis.

The low-volume hypothesis postulates that the loss of inhibition of epithelial sodium channels, because of *CFTR* dysfunction, leads to excess sodium and water reabsorption, resulting in dehydration of airway surface liquid. Concomitant loss of chloride efflux prevents the epithelium from correcting the low airway surface water volume. The subsequent decrease in periciliary water volume results in a reduction in the lubricating layer between epithelium and mucus, with compression of cilia by mucus causing inhibition of normal ciliary and cough clearance of mucus (Figure 1). According to this hypothesis, mucus on the epithelium forms plaques with hypoxic niches that can harbour bacteria, particularly *Pseudomonas aeruginosa*. This mechanism applies both in the airways and in the gastrointestinal and reproductive tracts.

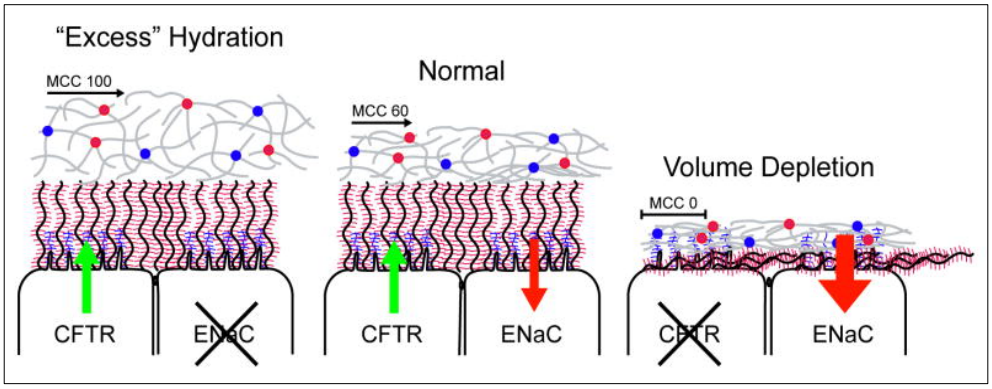

Figure 1. *CFTR* and ENaC channels influence water volume in ductal lumen. [8]

The high-salt hypothesis argues that in the absence of functional *CFTR*, excess sodium and chloride are retained in airway surface liquid. The increased concentration of chloride in the periciliary layer disrupts the function of important innate antibiotic molecules (i.e. human β-defensin 1, whose encoding gene is expressed at high levels in most surface epithelial cells throughout the lung [9]), allowing bacteria that are cleared by normal airways to persist in lungs. However, this hypothesis has

been already challenged by several reports showing no difference in Na+ hyperabsorption. [10]

Dysregulation of the host inflammatory response has been postulated as the putative basic defect in cystic fibrosis. Support for this hypothesis lies in the fact that abnormally high concentrations of inflammatory mediators are seen in cystic fibrosis cell cultures and uninfected ex-vivo tissue samples. Furthermore, findings from lung lavage studies show that inflammation is present in children as young as 4 weeks of age who are apparently free of infection. Levels of pro-inflammatory molecules such as interleukin 8, interleukin 6, tumour necrosis factor α, and arachidonic acid metabolites have been found increased in patients with cystic fibrosis. [11, 12] Apoptosis and activation of the nuclear factor-κB pathway also contribute to the self-perpetuating inflammatory cycle. [13] At the same time, concentrations of native anti-inflammatory substances such as interleukin 10, lipoxin, and docosahexaenoic acid are reduced, leading to an imbalance between pro-inflammatory and anti-inflammatory mediators, that favours unabated inflammation.

Another hypothesis suggests that primary pre-disposition to infection is a mechanism by which *CFTR* dysfunction leads to cystic fibrosis. In normal hosts, *P. aeruginosa* binds to functional *CFTR* and initiates an innate immune response, which is rapid and self-limiting. In patients with cystic fibrosis, an increase in asialo-GM1 in apical cell membranes allows increased binding of *P. aeruginosa* and *Staphylococcus aureus* to airway epithelium, without initiation of the *CFTR*-mediated immune response. The result is that in cystic fibrosis, the self-limiting response that quickly eliminates *P. aeruginosa* from the respiratory tract is lost at the same time as there is enhanced attachment of bacteria to the epithelial surface. [14]

Airway surface dehydration, due to an imbalance between *CFTR*-dependent Cl- secretion and ENaC-mediated Na+ absorption, and reduced ASL pH, due to impaired HCO3- secretion caused by *CFTR* dysfunction, are key pathogenetic mechanisms that trigger mucus plugging, chronic inflammation and impaired anti-bacterial host defence in CF airways. Intrinsic pH difference may exist between

CF and non CF-cells, and several reports give support to the hypothesis that a low extracellular pH might be a key factor in the susceptibility to infection in CF, although different and contradictory results were obtained so far. [10]

Signs and symptoms

CF manifestations are already apparent in utero and appear throughout life, encompassing many organ systems with great variability among patients. [1] The first problem encountered after birth is meconium ileus, which can present in as many as 20% of affected newborns. It is an obstructive condition secondary to inspissated material in the small and large bowels.

The lungs are normal in appearance at birth, but they shortly become infected and inflamed. Lung disease is in fact the most relevant in terms of symptoms, with recurrent cough, wheeze and pneumonia starting in infancy and continuing into adulthood, finally leading to pulmonary insufficiency, that is responsible for death in more than 80% of cases. Other respiratory symptoms, like nasal polyps and chronic sinusitis involve the upper airways.

In the past the main cause of death, occurring within months after birth, was malnutrition and failure to thrive. [1] Gastrointestinal disorders, in fact, were the first to be reported in conjunction with the cystic lesions of the pancreas that gave the disease its name. Despite improvements in disease management and the fact that airways disease is now the main focus of attention, gastrointestinal disorders continue to have a high impact on a patient's life. The spectrum and presentation of gastrointestinal disorders is very wide, even in patients with the same mutations. Distal intestinal obstructive syndrome, a condition analogous to meconium ileus, can present at any age, but it is especially common in adults. Impacted mucus and undigested food adhere to the small intestine mucosa and lead to acute abdominal distension and pain that must be differentiated with constipation, which can be a chronic condition.

Pancreatic insufficiency affects approximately 80% of patients, requiring pancreatic enzyme replacement therapy to avoid malnutrition, growth delay, vitamins and minerals deficiency and abdominal symptoms such as diarrhoea and bloating. Those patients who avoid pancreatic insufficiency still have a higher risk of developing acute pancreatitis.

Cystic fibrosis-related diabetes (CFRD) affects 2% of children, 19% of adolescents and 50% of adults. [15] Other gastrointestinal problems include gastroesophageal reflux, dysmotility, small intestine bacterial overgrowth and hepatobiliary disorders. [16]

Osteopenia can be detected in childhood as a consequence of vitamin D deficiency, poor nutritional status, physical inactivity, chronic inflammation and corticosteroid therapy. More than 50% of adults suffer from fragility fractures and kyphosis.

In the genitourinary system, sensitivity to *CFTR* alterations means that males are sterile due to azoospermia and congenital bilateral absence of *vasa deferentia* (CBAVD). Even carrier males can have CBAVD. Females are fertile, but often suffer from stress incontinence.

Diagnosis

Newborn screening for cystic fibrosis is offered in many European countries using the Guthrie heel prick test; diagnosis is then confirmed by the sweat test, the current gold standard procedure, [1] which measures the concentration of electrolytes in the sweat. Localised sweating is obtained by iontophoresis of the drug pilocarpine nitrate. In cystic fibrosis, sodium and chloride concentration will be above 59 mmol/L; normality is below 30 mmol/L. An intermediate concentration, between 30 and 59 mmol/L is often found in patients with milder *CFTR*-related diseases, but the diagnosis of classic cystic fibrosis is confirmed if two CF causing *CFTR* mutations are found. In infants, an intermediate concentration is less likely to be a true normal than in adults, and a borderline or normal sweat test with classic signs and symptoms is found only in 1–2% of patients.

Treatment

People with cystic fibrosis have complex care needs that demand specialist care delivered by a multidisciplinary team (including physicians, physiotherapists, psychologists, dieticians and social workers) adopting a global approach in specialised Centre. The life expectancy has increased significantly in successive patient birth cohorts as a result of more effective treatments and crucially because most patients attend CF Centres in line with the demonstration that patients who attend CF Centres for their care have better well-being and lung function than those who do not. Thus, the CF Centre has become the model of care for people with CF; patients should receive full care from the Centre or have local directed care supervised by the Centre. Therefore, standard care for cystic fibrosis requires multi-disciplinary teams and multi-professional approach.

The following professionals are reported as core staff for a CF Centre: respiratory paediatrician/pulmonologist, gastroenterologist, clinical microbiologist, medical support from trainee(s), clinical nurse specialist, specialist physiotherapist, specialist dietician, clinical psychologist, social worker, pharmacist, clinical geneticist, secretarial support and database coordinator. [17]

While there is no cure for cystic fibrosis, current treatment delays organ function decay, but it poses a high burden on the patient. Therapy includes respiratory care and infection prophylaxis, physiotherapy, nutritional and gastroenterologic care, management of diabetes, liver disease, low bone mineral density and arthropathy, male infertility, incontinence, psychological and psychiatric comorbidities. Among the new therapies available, there are currently four FDA-approved medications under the name of *CFTR* modulators, which basically restore and optimize the function of mutant *CFTR*, thus treating directly the underlying defect in CF: ivacaftor (Kalydeco® ≥4 months old), lumacaftor/ivacaftor (Orkambi®, ≥2 years old), tezacaftor/ivacaftor (Symdeko®, ≥12 years or older) and elexacaftor/tezacaftor/ivacaftor (Trikafta™, ≥12 years or older). Being mutation-specific,

these therapies are now available for patients who either carry a single copy of the *F508del* allele, two copies of *F508del* allele or for patients with some gating and conduction mutations.[18] The clinical trial of ivacaftor showed a reduction of sweat chloride concentration under the CF diagnostic range and an increase in lung function of 10%. [19, 20] The combination of ivacaftor with lumacaftor showed a modest clinical improvement as well, defined by a 40% reduction of the rate of pulmonary exacerbations, by an increase of body mass index (BMI) and, eventually, by a 4 to 7% increment of predicted FEV_1. [21] Symdeko® was associated with lower occurrence of pulmonary exacerbations and improvement of FEV_1 %pred, [22] however this combination did not affect BMI, which is usually correlated with better survival. So far, the triple combination therapy has demonstrated statistically and clinically meaningful results (i.e. +10-14% FEV_1 % pred), considering above all that this effect was achieved on patients who have at least one copy of the *F508del CFTR* mutation. [23] Whether these highly effective therapies will change the course of the disease and its burden remains an open question.

Prognosis

The median life expectancy for cystic fibrosis patients born in 2018 is 47.4 years, according to the North American CF patient registry, representing a huge improvement over the infantile death of the past that can be attributed to newborn screening and early and better therapy. [24] There is, however, great variability in the clinical course that translates into variability in prognosis.

Negative contributing factors include pancreatic insufficiency, female gender, *Pseudomonas aeruginosa*, *Staphylococcus aureus* and Burkholderia cepacia colonisation, predominance of respiratory symptoms and passive smoke exposure. Sex, genotype, newborn screening, meconium ileus and pancreatic status are defined as non-modifiable risk factors. The influence of these factors generally does not change over time, therefore they are considered time independent. Infection,

nutritional status, pulmonary exacerbation and respiratory illnesses vary over time and are the most commonly reported modifiable risk factors that influence CF lung function decline. Episodes of cough, fatigue, physical findings and other signs including rapid drop in lung function and weight loss define indeed pulmonary exacerbations (PE), which are associated with mortality [25], reduced quality of life, [26] unrecoverable loss of lung function [25, 27] and increased health costs.[28] Despite the lack of a globally accepted definition of PE, [29] any episode of clinical deterioration requiring antibiotic treatment seems to best characterizes such episodes.

As shown in Figure 2, lung disease decline varies according to different risk factors. FEV_1 is usually expressed as a percentage of predicted normal values, obtained from healthy populations of similar age, race and sex (for further details, see p. 20). Generally, FEV_1 % predicted observations are the highest and most stable between 5-10 years of age, afterwards the steepest decline is noted from 11 to 15 years of age followed by a reduced decline up to 30 years of age; after the age of 30 FEV_1 %predicted appears relatively stable.

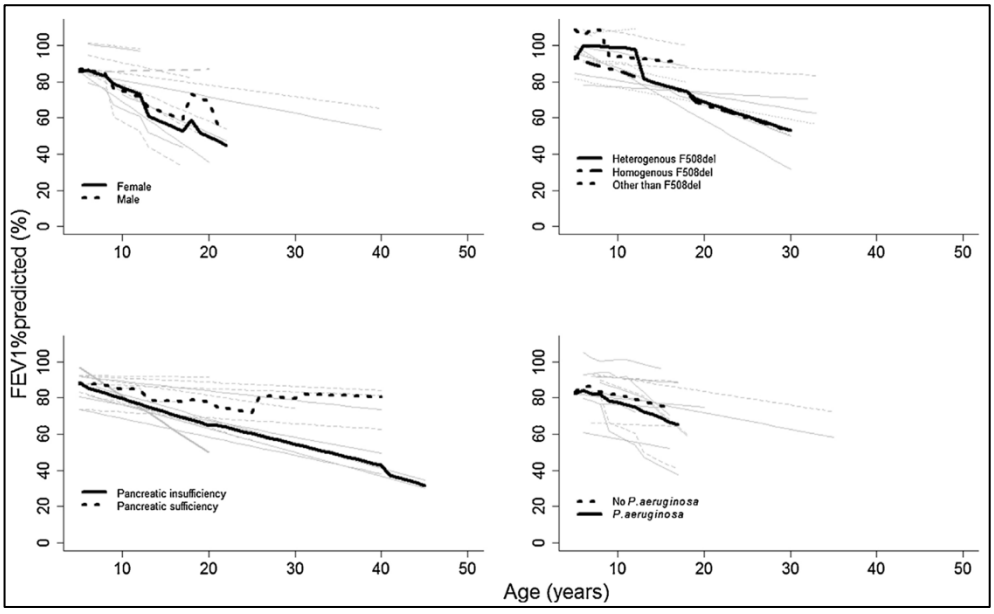

Figure 2. Lung disease progression in patients with CF. Weighted mean of FEV_1 %predicted over age for each risk factor is displayed in black. From reference [30]

Despite the change in FEV_1 %predicted is currently primarily described as a linear decline over a patient's lifetime, it is noteworthy that the relationship between FEV_1 %predicted decline and age is not static or directly proportional, but dynamic and time-varying, more specifically a nonlinear lung function progression over a patient's lifetime is frequently described in literature. [30]

As the median age of these patients increases - a UK model predicts that a child born with cystic fibrosis today will be 50 years of age or more [31] - an associated rise in gastrointestinal cancer is expected to occur as a consequence of both the primary disease and the use of immunosuppressive treatment following lung transplantation. [3]

Pulmonary Function Tests

Spirometry is conventionally used in the assessment of CF lung disease. FEV_1 reflects air flow obstruction in large airways only, whereas distal heterogeneous changes in airways structure may be masked, potentially with FEV_1 remaining within normal limits throughout childhood. Early changes in small airways are known to characterize cystic fibrosis pathophysiology [32] and early infection and inflammation, which occur soon after the diagnosis, have an impact on future health; therefore it becomes mandatory being able to recognize these alterations early in life. Nevertheless, spirometry requires collaboration, and it is pretty much under practised before 6 years of age, when relevant changes in small airways have already occurred.

The Multiple Breath Washout (MBW) is a gas washout technique, introduced in the '50, mainly used to quantify the *inhomogeneity* of gas distribution during tidal breathing, by examining inert gas clearance over a series of breaths. [33] Each MBW test consists of a washin and washout phase. In the case of exogenous inert gases (i.e., Sulfur hexafluoride: SF_6, helium: He, and methane), the gas is released at a known concentration during the washin phase and the washin phase is complete when the expired concentration reaches the concentration of the gas supplied. On the contrary, no formal washin phase is needed for endogenous gases (i.e., Nitrogen: N_2, and argon), and few tidal breaths

are performed to ensure that the concentration of N_2, for example, is stable. During the washout phase, the subjects inhale gases that do not contain the tracer gas, and the test is interrupted once the tracer gas concentration reaches a given fraction of its starting level.

Using MBW technique, distal lung units with slower time constants will contribute to the overall result obtained, by contributing to the expired nitrogen concentration at a later stage than faster emptying lung units. [34] In this way MBW is more reflective of peripheral airway function than FEV_1. Particularly, the lung clearance index (LCI) is the most commonly reported MBW index in current CF literature, and it is defined as the number of functional residual capacity (FRC) lung turnovers required to reduce alveolar tracer-gas concentration to a given fraction of its starting concentration, historically 1/40 (2.5%, $LCI_{2.5}$). It is calculated as the ratio of the cumulative expired volume (CEV) (total expired air volume to reduce the tracer gas at 1/40th of the starting concentration) to FRC; the latter is used as a denominator to correct for lung size, and leads to a standardized unitless index, [35] and may indicate some degree of lung hyperinflation. Thus, as lung ventilation worsens, the number of tidal breaths and expired volumes required to wash out the inert gas results in an increasing $LCI_{2.5}$ value. Healthy control $LCI_{2.5}$ values are between 5 and 7.5 in children age 6 and upwards into adulthood, [35] usually below 8.5 lung turnovers in healthy subjects. [36] Recently, normative values and upper limits of normal were generated for $LCI_{2.5}$ performed using an ultrasonic flowmeter (Exhalyzer® D). This limit was 7.91 for $LCI_{2.5}$ in Caucasian patients aged from 6 to 18 years old. [37]

Lung Clearance Index

$LCI_{2.5}$ derived from MBW, as a measure of ventilation inhomogeneity, has been shown to be a more sensitive measure of change in heterogeneous lung disease than FEV_1. There are only two prediction equations available in CF for $LCI_{2.5}$ so far. One is based on a small numbers of subjects (n=272) and explained the 44% of the variability of the response variable, under the linear model $\hat{Y}$=

6.25+0.66*age, [38] while the other only predicts $LCI_{2.5}$ for children under 6 years old of age. [39]. Recently, a multicentre 12-month longitudinal observational study in children aged 2.5-6 yrs was run at three North American CF Centres [40] with the main objective of defining the course of $LCI_{2.5}$ over time in health and disease. $LCI_{2.5}$ significantly deteriorated with a slope of 0.40 $LCI_{2.5}$ units/yr (95% CI, 0.14 to 0.66; P = 0.003) whereas in healthy children $LCI_{2.5}$ remained stable with age, with a slope of -0.04 $LCI_{2.5}$ units/yr (95%CI -0.12 to 0.04; P = 0.338).

When assessed longitudinally from 1978 to 1998 in 142 children, $LCI_{2.5}$ was shown to be the earliest and most predictive measure of disease progression. [41] High-resolution computed tomography (CT) allows visualization of structural changes along CF disease progression, carrying a radiation burden though. When compared to spirometry, $LCI_{2.5}$ proved to have better ability to detect structural lung abnormalities correlating with CT and reflected the extent of bronchiectasis and air trapping within CT scores. [35]

$LCI_{2.5}$ is sensitive to alterations in airway physiology throughout the airway tree, including very peripheral airway effects, and this probably accounts for its particular sensitivity in early CF. In adults with CF, abnormalities in $LCI_{2.5}$ appear to be caused by both increased dead space and increased specific ventilation heterogeneity. [42] $LCI_{2.5}$ as an outcome measure is still not fully used in clinical practice. However, $LCI_{2.5}$ has been shown to be superior to FEV_1 in detecting early disease, and it is now used as surrogate endpoint in clinical trials. However, the advantages may prove to be fewer in adults.

Patients with advanced CF lung disease have profound heterogeneity of ventilation, and both washin and washout are prolonged excessively. Testing in adults takes longer than in children, and a single washin/washout may take 10 minutes, which is then usually repeated in triplicate to derive an averaged value. From a purely practical point of view, it is therefore not appropriate to try assessing $LCI_{2.5}$ in those with severe CF. In subjects with the best-preserved spirometry, $LCI_{2.5}$ measurement

is easiest to perform, most reproducible and offers the most additional information. It thus seems sensible to restrict the measurement to those with FEV_1 impairment in the mild-moderate category (i.e., FEV_1 >40%). [20]

Changes in $LCI_{2.5}$ >25% can be considered physiologically relevant and greater than the biological variability of the test in cystic fibrosis. [43] $LCI_{2.5}$ has been shown to be a reproducible tool with low coefficients of variation within test occasions, between tests and between sites, the majority of which are between 3-7%. It is also shown to be highly discriminative in distinguishing between health and disease when comparing CF patients with healthy controls. [35]

Phenotyping CF

Generating clinical phenotypes facilitates assessment of disease prognosis, response to therapy, and provides insights into airway biology and disease pathophysiology. There are some studies available in CF that identify classes of patients (i.e. clusters), and the majority of these aimed to characterise the extent to which various *CFTR* alleles contribute to clinical variation in CF. These studies demonstrated that the degree of correlation between *CFTR* genotype and CF phenotype varies between its clinical components and is highest for the pancreatic status and lowest for pulmonary disease. The poor correlation between *CFTR* genotype and severity of lung disease strongly suggests an influence of environmental and secondary genetic factors (CF modifiers), which makes difficult clustering patients in well-defined clinical boundaries.

Recently, three distinct phenotypes of rapid decline were suggested, corresponding to early, middle, and late timing of maximal FEV_1 loss, in a large cohort 18.387 patients with CF. [44] The majority of variation among patient profiles was characterized by differences in mean longitudinal FEV_1 trajectories. Average degree of rapid decline was similar among phenotypes (roughly 23% predicted/yr); however, average timing differed, with early, middle, and late phenotypes experiencing rapid decline at 12.9, 16.3, and 18.5 years of age, respectively. Adult (n=211) clinical phenotypes

were identified using a clustering strategy that generated four and five phenotypes in another recent publication [45]. Each strategy identified 1) a low lung health scores phenotype, 2) a younger, well-nourished, male-dominated class, 3) various high lung health score phenotypes that varied in terms of age, sex and nutritional status. This multidimensional clinical phenotyping strategy identified classes with expected microbiology results and low risk clinical phenotypes with pancreatic sufficiency. Eventually, in a younger cohort of patients with CF (n=212) used by other authors [46] to generate a scoring system in CF, it was shown that variables that contribute to predicting severity did not vary with gender. To the best of my knowledge, only one study used $LCI_{2.5}$ to define physiological phenotypes in 80 school children with different lung diseases, including CF. [47] Three main phenotypes were identified, and the third one was characterized by increased global and diffusion- and convection-dependent ventilation inhomogeneity as well by the highest proportion and frequency of exacerbations and hospitalization.

Hypothesis and Objective

There are several underlying hypotheses contributing to the development of this project. Given the ascending role of $LCI_{2.5}$ as early marker of lung disease, it is expected that the association of $LCI_{2.5}$ with the variables currently used in clinic to assess CF disease, including well known prognostic factors, could be different from that of FEV_1, thus characterizing clinical phenotypes based on early signs of lung disease. It is therefore expected that different phenotypes will describe different trajectories of lung function over time. Eventually, if $LCI_{2.5}$ tracks lung disease better than FEV_1 alone, association of $LCI_{2.5}$ with time to pulmonary exacerbation should be present as well.

Considering the experience collected from 2014 by the Regional CF Centre on MBW and the empirical-driven use of $LCI_{2.5}$ in the clinical decision making, the main objective of this work is to ascertain if $LCI_{2.5}$ could be worth to be implemented in a pediatric CF clinic, as a routine respiratory assessment complementary to FEV_1. Understanding the clinical behaviour of $LCI_{2.5}$ under multiple clinical perspectives will add to the current literature novel information about the clinical utility of MBW for children with CF.

Methods

According to the current clinical practice, patients attending the CF Center of Milan perform $LCI_{2.5}$ as part of their clinical review mostly during outpatient follow-up visits since 2014. Spirometry and $LCI_{2.5}$ were performed according to the available guideline [48, 49] and standard operating procedure. [33, 50] Written, informed consent was obtained from patients prior to each annual review. Data were collected between October 2014 and September 2019. For the present analyses, only spirometry and $MBWN_2$ test results from children in clinical stable condition where considered, as defined by absence of change in treatment (i.e., antibiotics treatment, systemic corticosteroids), hospital admissions and/or signs and symptoms of pulmonary exacerbation.

Spirometry

Spirometry requires a full inspiration followed by a forced expiration and it measures among other indices the FEV_1, which is an indicator of airways obstruction. Performing spirometry requires the patient's cooperation, meaning that it cannot be easily accomplished below 3-6 years of age.

Forced expiration manoeuvre and vital capacity manoeuvre were measured in this study using a standardized method: the absolute values of FVC and FEV_1 were measured and expressed as a percentage values for age, height and sex, and as Z-score, according to Quanjer's equation developed under the Global Lung Function (GLI) initiative. [51] Lung volumes were measured by a 830 litres plethysmograph (Master Screen Body 4.2, E. Jaeger GmbH) in the sitting position, according to ATS/ERS guidelines. [48] Flow and volume were measured by a pneumotachograph with a 0.036kPa L/-1s resistance and 160ml dead volume. Calibration with three litres syringe cycled three times was performed every day before testing. Patients with CF are well trained in performing spirometry as it is a routine assessment during their follow-up visits. However, operators conducting the test always met the minimum technical standards, as recently updated. [49]

Multiple breath washout

The MBW test measures the distribution homogeneity of an inert gas, i.e. nitrogen (N_2). Each test consists of a washin and a washout phase. During washin, the tracer gas is administered at a known concentration until the expired concentration equals the delivered one. During washout the subject breathes a gas mixture that does not contain the tracer gas. The tracer gas concentration decreases exponentially due to the differences in lung branches generations, and the washout phase ends when the concentration reaches the historical value of 1/40th of the initial concentration. [52] Several MBW devices are currently available on the market and most system have a sidestream gas analyzer, which can be a respiratory mass spectrometer, an N_2 analyzer or an infrared gas analyzer. When nitrogen is used as tracer gas ($MBWN_2$), no formal washin phase is required, since N_2 is an intrinsic gas, and it is sufficient to ensure that its concentration in the lungs is stable; for washout, 100% oxygen is usually used. The MBW test requires only a tight facemask or mouthpiece and quiet tidal breathing for 2 to 10 minutes per test, and it is usually performed in triplicate in a sitting position. $MBWN_2$ performed via Exhalyzer® D (Ecomedics AG, CH) is the modality chosen by the European Cystic Fibrosis Society and the one most frequently used in Italy. [53]

Lung clearance index

An open-circuit MBW hard- and software package with nitrogen as tracer gas ($MBWN_2$) was used (Exhalyzer® D and Spiroware 3.2.2 Ecomedics AG, CH) and calibration and measurement procedures were performed as suggested. [33, 50] Only results from three reproducible runs were considered, defined as a variation of FRC and $LCI_{2.5}$ values within 10%. $LCI_{2.5}$ together with $Scond^{*VT}$ and $Sacin^{*VT}$, believed to reflect convective gas mixing in the conducting airways and considered to best represent diffusion–convection interaction within the acinus respectively, were therefore recorded, before patients performed spirometry.

In children under 6 years old, ULN was calculated based on the equation reported by Lum et al. [39] For the paediatric age, the cut-off of normality equals of 7.91 was adopted, as recently published. [37] An adequate environment with adequate distraction for younger children was assured during each test. [54] Data from earlier software versions were re-run in the version of Spiroware 3.2.2.

Statistical Analysis

Data were collected for each patient's visit from 2014 up to September 2019. Four main objectives were considered during the analysis, adopting either a cross-sectional or a longitudinal design. Particularly, the aims were: investigating the association between lung function measures and selected variables; profiling paediatric patients according to their clinical characteristics in order to generate plausible phenotypes, and subsequently evaluating the longitudinal association between lung function measures and selected variables between phenotypes. The last objective is investigating the association between selected variables and the risk of pulmonary exacerbation recurrence over the whole follow-up period.

To comply with the first objective, data from paediatric subjects with at least one $MBWN_2$ test performed during an outpatient visit were used, whereas the subset of paediatric patients who performed at least two $MBWN_2$ tests during regular outpatient visits was considered for the subsequent objectives. Specifically, data collected on the first evaluation were used to define clusters while follow-up data were used to fit longitudinal models to catch variation of $LCI_{2.5}$ and FEV_1 over time and to model pulmonary exacerbation as recurrent time-to-event data.

Before any analysis, the whole database was reviewed for errors, incomplete or missing information. Variables were categorized mainly for descriptive purposes according to the commonly accepted criteria reported in CF literature. To adjust for differences in growth and development (age, sex and height), lung function results evaluating airflow obstruction (FEV_1) were expressed also as Z-scores. $LCI_{2.5}$ as well was reported descriptively as Z-score, using the newly normative data published by

Anagnostopoulou et al. [37] Descriptive analysis was reported according to the number of $MBWN_2$ tests performed for each individual. Continuous variables were reported as mean and standard deviation, and as median together with first and third quartile; categorical variables were summarized as counts and percentage.

Association between lung function measures and selected variables.

Ordinary least-square (OLS) regression models were used to fit the relationship between FEV_1 %predicted, the dependent variable, and age, sex, pancreatic status, presence of diabetes, genotype, *Pseudomonas aeruginosa* infection and nutritional status as independent variables; a distinct linear regression model was used to assess the association of the same independent variables with $LCI_{2.5}$. In order to identify the best age-specific relationship, several models with a different number of spline knots were fitted on the dataset before imputation, using 3 to 6 knots, under the hypothesis that lung function and age were best summarized by a non-linear relationship indeed. All the models with a different number of spline knots were compared using Akaike information criterion (AIC), as a mean for model selection. Among the candidate models, the preferred model was the one with the minimum AIC value. Once identified the best probabilistic model, since $LCI_{2.5}$ (right skewed) and FEV_1 %predicted (left skewed) departed from normality, the Box-Cox method was implemented to suggest a transformation of these dependent variables when regressed as a combination of the independent variables (sex, pancreatic status, presence of diabetes, genotype, *Pseudomonas aeruginosa* infection and BMI).

Final models were therefore generated after multiple imputation (MI), since some data regarding *Pseudomonas aeruginosa* infection, airflow obstruction (FEV_1 %predicted) and indices of ventilation inhomogeneity were missing. Some missing indices derived from $MBWN_2$ were undetectable by the software machine, namely $Sacin^{*VT}$ and $Scond^{*VT}$, however, these were not considered predictors in

regression analyses. The distribution of the variables of interest according to whether FEV_1 or PSA were missing or not was explored by visual inspection of histograms.

Multiple Imputation (MI) was performed through the MICE algorithm available in the mice package (version 3.7.0). It allows for the specification of a separate imputation model for each variable in the dataset and it is implemented as a series of univariate equations. Variables values were imputed using the following built-in imputation models: predictive mean matching (a semi-parametric method which limits imputations to observed values only) for FEV_1 and proportional odds model for *Pseudomonas aeruginosa*, considered as an ordinal variable (free, intermittent and chronic). The wrapper function with.mids() from the mice package was used to obtain pool estimates of imputed datasets in one step.

The variability of the multiple imputation (MI) estimates consists of two components: variability within imputations and variability between imputations. The precision of the MI estimates depends not only on the number of observations in the sample, but also by the number of imputations (*m*). Therefore, several imputations were created, adopting the rule of thumb to use whatever the average percentage rate of missingness is, considering however that substantive conclusions are unlikely to change as a result of raising *m* beyond *m*=5, from an accuracy point of view [55].

The pool() function uses Rubin's combination rules to obtain the estimates from multiply imputed data. [56] It is assumed that complete data inferences about the population parameter of interest (Q) are based on the normal approximation $\mathrm{Q} - \widehat{\mathrm{Q}} \sim \mathrm{N}(0, \mathrm{U})$, where $\widehat{\mathrm{Q}}$ is a complete data estimate of Q and U is the associated variance for $\widehat{\mathrm{Q}}$. With missing data, estimates of the parameters of interest are calculated on each of the m imputed datasets to give $\widehat{\mathrm{Q}}_{1,\ldots,}\widehat{\mathrm{Q}}_{\mathrm{m}}$ with associated variances U_1, …, U_m. For a single regression coefficient, the MI overall point of estimate is the average of the *m* estimates of Q from the imputed datasets, $\bar{Q} = \frac{1}{m}\sum_{i=1}^{m} \widehat{\mathrm{Q}}_1$. The associated total variance for his overall MI

estimate is $T = \bar{U} + \left(1 + \frac{1}{m}\right)B$, where $\bar{U} = \frac{1}{m}\sum_{i=1}^{m} \hat{U}_1$ is the estimated within imputation variance and $B = \frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{Q}_1 - \bar{Q}\right)^2$ is the between imputation variance. When B dominates $\bar{U}$ greater efficiency, and then more accurate estimates, can be obtained by increasing *m*; on the contrary, when $\bar{U}$ dominates B, little is gained from increasing *m*. [56]

Diagnostic checks were run to ensure that imputed values were plausible, describing the relative increase $r = \frac{(1+m^{-1})B}{\bar{U}}$ in the variance (RIV) due to multiple imputations as well. Relative efficiency, defined as $RE = \frac{1}{1+\frac{FMI}{m}}$, was also calculated in order to give information about the precision of the parameter estimate. Fraction of missing information (FMI) was derived as the ratio between $RIV + \frac{2}{df+3}$ and $1 + RIV$,where *df* is the degree of freedom of pooled results.

Finally, the function mi.anova() from miceadds package (version 3.8-9) in R was used to obtain results of the Wald test.

Goodness of fit and model assumptions were judged from inspection of normal Q-Q plots, the distribution of residuals and from density plots of residuals; Shapiro-Francia test for normality and Breusch–Pagan test were used as well to check for heteroskedasticity.

Profiling paediatric patients according to their clinical characteristics.

Clustering was the approach selected to generate phenotypes based on the variables collected throughout the study. In adjunct to age, sex and *CFTR* genotype, clustering was based also on nutritional status, i.e. BMI Z-score and pancreatic exocrine function, on the presence of CFRD, and on colonization by *Pseudomonas aeruginosa.* To characterize lung disease, indices derived by $MBWN_2$ test were used, namely LCI, $Sacin^{*VT}$ and $Scond^{*VT}$. The number of pulmonary exacerbations in the twelve months preceding LCI together with the number of hospitalizations in the

previous year were also used in order to define clusters not only based just on classic markers of the disease. Clustering tendency of the data was assessed by measuring the probability that a given dataset is generated by a uniform data distribution (*Hopkins* statistic).

Given the mixed-data type defining the current dataset, distance between pairs of observations was obtained using Gower's metric, available with the function daisy(), which use the general dissimilarity coefficient of Gower [57]. Indeed, presence of CFRD, *CFTR* genotype grouped according their alleles distribution, pancreatic exocrine function, and presence of *Pseudomonas aeruginosa* infection are categorical variables while lung function, age, nutritional status and pulmonary exacerbations or hospitalizations are continuous. By the means of Gower's index, two individuals *i* and *j* can be compared on a character *k* and assigned a score $s_{ij,k}$, zero when *i* and *j* are considered different and a positive fraction, or unity, when they have some degree of similarity. In case of dichotomous variables, a character can be non-existent in both *i* and *j*. Therefore, the possibility of making comparison can be represented by a quantity $\delta_{ij,k}$, equal to 1 when character *k* can be compared for i and j, zero otherwise. When $\delta_{ij,k} = 0$, $s_{ij,k}$ is set to zero. The similarity between *i* and *j* is defined as the average score taken over all possible comparisons: $S_{ij} = \sum_{k=1}^{v} s_{ij,k} \Big/ \sum_{k=1}^{v} \delta_{ij,k}$. When $\delta_{ij,k} = 0$ for all characters, S_{ij} is undefined. When all comparisons are possible, $\sum_{k=1}^{v} \delta_{ij,k} = v$, the total number of characters; otherwise it is the number of characters over which the comparison is made. For dichotomous characters the presence of the character is denoted by + and its absence by -. When there are no unknown values of character *k*, four different combinations of its values may occur for two individuals. For qualitative characters, $s_{ij,k} = 1$ if the two individuals *i* and *j* agree in the k^{th} character and $s_{ij,k} = 0$ if they differ. For quantitative characters with values $x_1, x_2, \ldots, x_n$ of character *k* for the

total sample of n individuals, $s_{ij,k} = 1 - \left|x_i - x_j\right| / R_k$. Here R_k is the range of character *k* and may be the total range in the population or the range in the sample. When $x_i = x_j$ then $s_{ij,k} = 1$, and when x_i and x_j are at opposite ends of their range, $s_{ij,k}$ is a minimum (0 when R_k is determined from the sample). With intermediate values, $s_{ij,k}$ is a positive fraction.

The use of Gower's distance limited the possible algorithms to be explored in order to identify clusters within this CF cohort, and therefore the final choice was made between two *bottom-up* hierarchical methods, i.e. AGglomerative NESting (agnes) and Hierarchical CLUSTering (hclust), and one partitioning method, i.e. Partitioning Around Medoids (PAM), an approach which is based on medoids, a robust alternative to k-means.

The first two methods mentioned above are used to group objects based on their similarity. The algorithm starts by treating each object as a single-element cluster, and pairs of clusters are successively merged until all clusters have been grouped into one big cluster containing all objects. In this sense, this approach defines this method as *bottom-up*. Finally, objects and/or clusters in close proximity are linked together via the linkage function, which takes the distance information and groups pairs of objects into clusters to create bigger ones. For hierarchical methods, the best linkage function was selected comparing correlation coefficients between cophenetic and the original distance, resulting from different clustering methods: average, single, complete and Ward.

PAM classifies observations into multiple groups based on their similarity as well, but it is defined as a partitioning method, given that it ends with a division of the set of data objects into non-overlapping subsets (clusters). Each cluster, whose number must be specified a priori, is represented by one of the objects in the cluster. This object is called *medoid* and it is the most centrally located point in the cluster and can be considered representative of the members of that cluster; in this sense,

PAM is considered to be a more robust approach in the presence of outliers. In practice, after finding a set of *k* medoids, clusters are constructed by assigning each observation to the nearest medoid.

To determine the optimal number of clusters, *elbow, silhouette* and *gap statistics* were taken into consideration, together with Frey and *C*-index [58] and the visual inspection of cluster plots, whereas to select the best clustering algorithm, several methods for clustering validations were implemented and compared.

As internal cluster validation measures, whose purpose is to describe the goodness of clustering structure without reference to external information, *silhouette coefficient* and *Dunn* index were used. The *silhouette* analysis measures how well an observation is clustered and it estimates the average distance between clusters. [59] For this study, the overall average silhouette width, which is the average of the silhouettes for all objects *i* in the whole dataset, was used as a means to compare clustering solutions. Larger values suggest stronger clustering structure. The *Dunn* index computes the distance between the objects in the same cluster as measure of intra-cluster compactness. Therefore, Dunn index is expected to be maximized if data contains compact and well-separated clusters.

Furthermore, few other indices were evaluated through the function `cluster.stats()` available in the `fpc` package. These were the *average between* and *within* distance between clusters.

Eventually, consistency measures of clustering results were obtained by adopting the `clValid` package, which reports the average proportion of non-overlap (APN), the average distance (AD), the average distance between means (ADM) and the figure of merit (FOM). Smaller values correspond to highly consistent clustering results. Overall, these are considered stability measures and guided the choice towards the best clustering algorithm.

Longitudinal association between lung function measures and selected variables between phenotypes.

The clustering approach allowed me to generate distinct phenotypes, on which longitudinal variation of $LCI_{2.5}$ was further explored. A mixed effects modelling was used taking into consideration the correlation between subjects as the only random effect, whereas cluster and time were considered fixed effects. Being interested in the difference between phenotypes, interaction between cluster and time was also considered. Linear mixed-effects models were fitted using the R package nlme (version 3.1) using lme() function.

As regards to $LCI_{2.5}$, to account for the variance heterogeneity found between the clusters, two models were evaluated, one with a general correlation structure and another one with *varIdent*(.) variance function. A REstricted Maximum Likelihood (REML)-based Likelihood Ratio (LR) test was used to guide the selecting process towards the best fitting model. Models with constant variance and model with different variance in each cluster were therefore compared with the use of the anova () function. Goodness of fit was evaluated by the means of separate scatterplots of Pearson residuals and Q-Q plots per cluster, in order to remove the influence of any residual correlation. The same approach was considered for modelling FEV_1 %predicted longitudinally. For both outcome measures, autocorrelation was checked by the means of ACF() by the same nlme package.

Risk factor for pulmonary exacerbation as recurrent time-to-event data. In order to take into account the portion of the pulmonary exacerbations information collected in paediatric patients, extensions of the Cox proportional hazards model were investigated to model recurrent event data.

The standard Cox proportional hazard model specifies the hazard of i^{th} individual as:

$$\lambda(t) = \lambda_0(t) \exp(\beta x_i)$$

where $\lambda_0(t)$ is an unspecified baseline hazard function and β is the vector of regression coefficients, x_i is the vector of covariates of the ith individual. Due to the independence assumption, this model is only appropriate for modelling the time to first event, therefore causing a loss of many information available on the cohort when recurrent events are there.

The conditional model of Prentice, Williams and Peterson offers a viable solution to handle multiple and ordered events [60], such as pulmonary exacerbations in cystic fibrosis. Under this model, an individual is not at risk for the kth event until he/she has experienced event k-1st. Time intervals are defined as: (entry time, first event], (first event, second event],…,(m^{th} event, last follow-up], but each event is assigned to a separate stratum (i.e. pulmonary exacerbation). The use of time-depending strata means that the underlying hazard function may vary from event to event, being specified for the jth event for ith individual as:

$$\lambda(t) = Y_{ij}(t)\lambda_{0j}(t)exp\big(X_i(t)\beta_j\big)$$

where the at-risk indicator, $Y_{ij}(t)$, is defined as zero until the j-1st event and only then becomes one. Once the jth event occurs, $Y_{ij}(t)$ becomes 0 again. Overall, this model takes into account the within-subject correlation due to event dependency (i.e. baseline risk for a second PE may be higher in CF than baseline risk of first PE) and within-subject correlation due to heterogeneity, being suitable for analysing this type of recurrent events. Risk group interactions was assessed by the means of LR-test.

All paediatric patients who performed at least one $MBWN_2$ test during regular follow-up visits were considered for this final analysis. An event (pulmonary exacerbation, PE) was defined as children requiring hospitalization or antibiotic administration. Individuals were censored when they were still in follow-up at the time of the present data collection; when patients experienced PE, the event was coded as recurrence and follow-up was ended from that moment onwards. Besides $LCI_{2.5}$, other covariates included in the model are children's BMI Z-score, sex, age, pancreatic status, *CFTR* genotype and *Pseudomonas aeruginosa* colonization at the first $MBWN_2$ test. *Pseudomonas*

aeruginosa was categorized into two groups (chronic *versus* intermittent-free), as well as *CFTR* genotype (*F508del* homozygotes *versus F508del* heterozygotes or other mutations). Because of collinearity, FEV_1 could not be included into the same model with $LCI_{2.5}$, and two adjusted models were fitted, one with $LCI_{2.5}$ and another one with FEV_1.

Prentice, Williams and Peterson-gap time risk intervals were used, since this allows for predictions from the time of the previous PE, thus allowing to explore the overall association between pulmonary exacerbation episodes and each covariate. Analyses were performed using packages *survival, survminer* and *rms* available in R. [61]

Results

Study population

Between October 2014 and September 2019, 939 $MBWN_2$ tests were performed on 412 patients with cystic fibrosis. One hundred thirty-one patients performed the $MBWN_2$ once, 121 two tests in separate occasions, 96 patients performed three tests, 46 patients were followed up four-times and 16 patients had the $MBWN_2$ performed 5 times; 76.8% of $MBWN_2$ tests were done during outpatient visits. In line with the purposes of the present discussion, 218 $MBWN_2$ tests performed during hospital admissions have been excluded from the analysis, and only results from children in clinically stable conditions while attending outpatient clinic have been considered.

Table 1 shows the characteristics at baseline of paediatric patients who performed at least one $MBWN_2$ test during regular follow-up visits only (n=245).

Table 1. Sample characteristics

Subjects, *nr*	245
Male/Female, *nr(%)*	136 (55.5) /109 (44.5)
Age, *yrs*	11.9 (3.4)
Age, *min-max*	5 – 17.9
BMI, Z-score	-0.5 (0.9)
F508del/*F508del*, *nr (%)*	50 (20.4)
F508del/ other, *nr (%)*	105 (42.9)
other/other, *nr (%)*	90 (36.7)
CFRD, *nr (%)*	4 (2)
Pancreatic Insufficiency, *nr (%)*	141 (57.6)
Pseudomonas aeurginosa chronic infection*, *nr (%)*	48 (19.6)

*Values are expressed as absolute number(percentage); *17 microbiological data are missing*

CFTR genotype, diabetes, pancreatic insufficiency and infection by *Pseudomonas aeruginosa* (PSA) are the considered prognostic negative factors. Overall, few children show unfavourable prognostic

factors in large percentages. Considering the young age of the present cohort, as expected few children have diabetes and in a very small percentage chronic infection by *Pseudomonas aeruginosa* was detected; this percentage is slightly above the national mean percentage (≈15%). [62]

Table 2 reports the anthropometric and pulmonary functions of this subsample. The high mean FEV_1 %predicted reflects the young age of patients attending this mostly paediatric CF centre: lung function interquartile range (IQR) for these children is 87.9-110.7%. Particularly, just 2 patients (1%) express a severe lung disease; the majority (89.8%) has a normal/mild lung disease, defined as FEV_1 %predicted equal or above 70%. By adopting the lower limit of normal (LLN) based on the 5th centile of FEV_1 (-1.64 Z-score) [51], the percentage of patients considered *normal* drops slightly (83.1%).

Table 2. Pulmonary characteristics	
Subjects, *nr*	245
FEV_1, *% predicted*	98.4 (19.2)
FEV_1, Z-score	-0.1 (1.6)
FVC, *% predicted*	105.9 (16.1)
FVC, Z-score	0.5 (1.4)
$LCI_{2.5}$	9.9 (3.3)
$LCI_{2.5}$, CV%	4.2 (2.3)
$LCI_{2.5}$, Z-score[¶]	4.3 (8.58)
Scond[*VT]	0.062 (0.033)
Scond, *CV%*	31.7 (28.5)
Sacin[*VT]	0.141 (0.111)
Sacin, *CV%*	31.0 (26.0)

Spirometry data belong to 225 patients. Sacin was not calculated in 7 children. All displayed values are expressed as mean (standard deviation). ¶ median with interquartile range (IQR).

The clinical evaluation of lung health changes dramatically if one implements the information derived by the $MBWN_2$ test. First $MBWN_2$ was performed on a 5 years old child. Fifty percent of paediatric

subjects performed this lung function test between 9 and 15 years of age, with 25% of the highest values ranging from 11.2 and 23.4. Compared to information derived by the analysis of FEV_1 %predicted alone, percentage of children considered with *normal* lung disease drops down to 35.1%, that is when $LCI_{2.5}$ values are below ULN, according to patients' age (*see* pg. 22).

The clinical advantage of having the $MBWN_2$ test available in the clinic is that it can be used complementary to classical spirometry results. Indeed, information derived by the intersection of the identified cut-offs (*see* pg. 22) as shown in Figure 3 is interesting. If patients in quadrant 1 (Q1) and quadrant 2 (Q2) can be generally identified as those with a normal/mildly abnormal pulmonary function as detected by FEV_1 alone ($FEV_1 \geq 70\%$ predicted), the $MBWN_2$ test adds the information that only patients in Q1 can be considered really healthy, from a lung perspective. As far as quadrant 3 (Q3) is concerned, the clinical meaning of the two lung function variables is here in agreement, showing a moderate-severe lung function (FEV_1) and a very inhomogeneous ventilation of the lungs ($LCI_{2.5}$). Despite it is unusual that subjects with an impaired airflow passage could have a homogenous gas mixing, one patient is represented in Quadrant 4 (Q4) when adopting FEV_1% predicted and two when adopting FEV_1 Z-score.

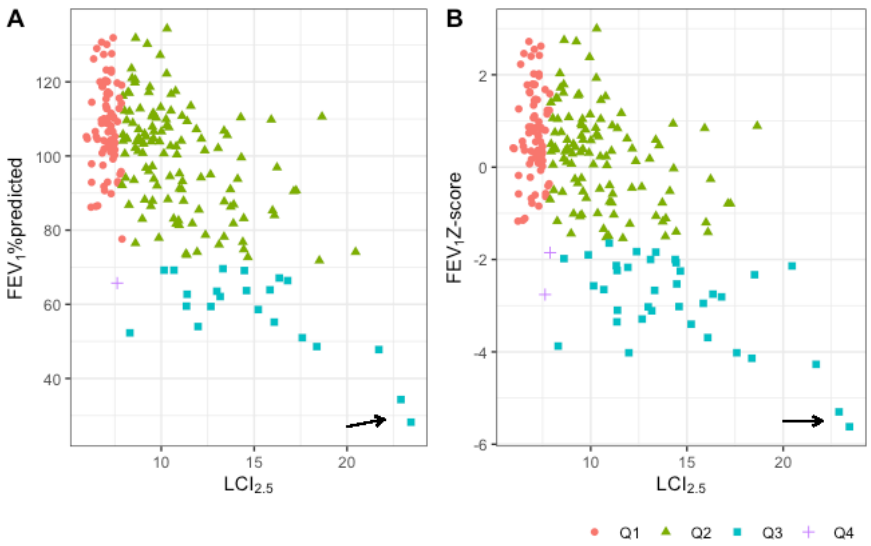

Figure 3. Patients are grouped (Q, Quadrant) according to their LLN FEV_1 and $LCI_{2.5}$ values. Black arrows identify extreme observations.

Table 3 reports the clinical characteristics of children according to the first three quadrants, taking as reference the plot B in Figure 3, which adopts Z-score of FEV_1, and thus takes into account the differences in growth and development (age, sex and height). The two patients with poor measures of airflow obstruction but normal $LCI_{2.5}$ are not reported in table 3. However, these patients were not diabetic, one was chronically infected by *Pseudomonas aeruginosa* and both were pancreatic insufficient.

Table 3. Sample characteristics stratified by air flow obstruction (FEV_1 LLN) and ventilation inhomogeneity ($LCI_{2.5}$) severity

	Quadrant 1 (n=79)	*Quadrant 2 (n=108)*	*Quadrant 3 (n=36)*
Age, yrs - *mean(SD)*	11.6 (3.4)	11.8 (3.3)	12.6 (3.6)
Age, yrs - *median(Q1;Q3)*	11.4 (8.8; 14.2)	11.6 (9.1; 14.6)	12.4 (9.5; 15.7)
Male/Female, nr	39/40	63/45	21/15
F508del/*F508del*, *nr(%)*	6 (7.6)	32 (29.6)	6 (16.7)
F508del/ other, *nr(%)*	30 (38.0)	50 (46.3)	17 (47.2)
other/other, *nr(%)*	43 (54.4)	26 (24.1)	13 (36.1)
CFRD, *nr(%)*	1 (2)	2 (2)	1 (3)
Pancreatic Insufficiency, *nr(%)*	19 (24.1)	82 (75.2)	29 (80.6)
Pseudomonas aeurginosa chronic infection*, *nr(%)*	6 (7.6)	22 (20.4)	17 (47.2)
BMI, Z-score - *mean(SD)*	-0.3 (1.0)	-0.6 (0.8)	-1.0 (0.8)
BMI, Z-score - *median(Q1;Q3)*	-0.2 (-0.8; 0.3)	-0.5 (-1.1; -0.1)	-1.2 (-1.6; -0.5)
FEV_1, % predicted - *mean(SD)*	108.6 (11.2)	102.7 (12.1)	64.9 (12.3)
FEV_1, % predicted - *median(Q1;Q3)*	108.8 (101.6; 116.8)	104.0 (93.4; 110.5)	68.1 (59.2; 74.1)
FEV_1, Z-score - *mean(SD)*	0.7 (1.0)	0.2 (1.0)	-2.9 (1.0)
FEV_1, Z-score - *median(Q1;Q3)*	0.7 (0.1; 1.4)	0.3 (-0.6; 0.9)	-2.7 (-3.4; -2.1)
FVC, % predicted - *mean(SD)*	111.9 (11.6)	109.3 (13.0)	83.4 (14.1)
FVC, % predicted - *median(Q1;Q3)*	111.4 (103.3; 120.2)	111.0 (99.9; 118.9)	85.2 (72.5; 91.8)
FVC, Z-score - *mean(SD)*	1.0 (1.0)	0.8 (1.1)	-1.4 (1.2)
FVC, Z-score - *median(Q1;Q3)*	1.0 (0.3; 1.7)	0.9 (0.0; 1.6)	-1.3 (-2.4; -0.7)
$LCI_{2.5}$ - *mean(SD)*	7.08 (0.44)	10.62 (2.37)	14.30 (3.78)
$LCI_{2.5}$- *median(Q1;Q3)*	7.10 (6.85; 7.39)	9.96 (8.85; 11.44)	13.35 (11.38; 16.15)
$LCI_{2.5}$ CV% - *mean(SD)*	3.9 (2.1)	4.3 (2.4)	4.5 (2.4)

$LCI_{2.5}$ CV% - *median(Q1;Q3)*	3.4 (2.2; 5.5)	4.0 (2.5; 5.7)	4.2 (2.4; 6.5)
$Scond^{*VT}$ - *mean(SD)*	0.036 (0.032)	0.073 (0.025)	0.080 (0.022)
$Scond^{*VT}$- *median(Q1;Q3)*	0.03 (0.022; 0.041)	0.071 (0.056; 0.091)	0.079 (0.070; 0.090)
$Scond^{*VT}$ CV% - *mean(SD)*	47.1 (32.7)	24.6 (23.1)	15.2 (9.8)
$Scond^{*VT}$ CV% - *median(Q1;Q3)*	40.0 (25.1; 59.6)	16.4 (9.1; 29.1)	13.5 (6.8; 21.8)
$Sacin^{*VT}$ - *mean(SD)*	0.088 (0.066)	0.144 (0.100)	0.246 (0.153)
$Sacin^{*VT}$- *median(Q1;Q3)*	0.07 (0.046; 0.106)	0.119 (0.082; 0.181)	0.231 (0.122; 0.293)
$Sacin^{*VT}$ CV% - *mean(SD)*	39.0 (27.5)	28.1 (24.9)	21.4 (23.2)
$Sacin^{*VT}$ CV% - *median(Q1;Q3)*	32.6 (18.4; 54.3)	18.4 (10.7; 40.7)	11.1 (6.4; 26.2)

If not specified, displayed values are expressed as absolute number(percentage). Q1 and Q3 denote first and third quartiles. Data in Quadrant 1 for Sacin belong to 78 patients; data in Quadrant 2 for Sacin belong to 105 patients and to 35 patients in Quadrant 3.

It is of clinical interest that low prevalence of *Pseudomonas aeruginosa*, a low degree of ventilation inhomogeneity ($LCI_{2.5}$, $Sacin^{*VT}$ and $Scond^{*VT}$), a low percentage of patients with pancreatic insufficiency and with a genotype mainly represented by one variant only, i.e., *F508del*, are the main clinical features which define at best children in the first quadrant. As far as FEV_1 %predicted is concerned, it poorly discriminates subjects between quadrant 1 and quadrant 2 whereas $LCI_{2.5}$ does.

The relationship between $LCI_{2.5}$ and FEV_1 is represented in Figure 3. FEV_1 is inversely correlated with $LCI_{2.5}$, whether expressed as FEV_1% predicted (r=-0.64 P<0.0001) or FEV_1 Z-score (r=-0.63 P<0.0001). When the two most extreme observations are removed, the strength of such relationship is slightly reduced (r=-0.58 P<0.0001). These two observations belong to two males of 17.4 and of 14.4 years old with FEV_1 below 30 %predicted.

As far as the relationships of lung function with age are concerned (Figure 4), $LCI_{2.5}$ and age are well represented by a linear function with a correlation equal to r=0.16 (P=0.011). By the contrary, the relationship between age and indices of airflow obstruction, either FEV_1% predicted (r=-0.15, P=0.029) or FEV_1 Z-score (r=-0.14, P=0.03), is better represented by non-linear function. Restricted

cubic splines were used indeed to explore the decrease/increase of lung function over age, although GLI equations already demonstrated a near-linear decline in the FEV_1 throughout childhood without acceleration or deceleration during adolescence. [63] It is reported that $LCI_{2.5}$ has an inverse relationship with height until adolescence, [54] here not detectable. Interestingly, taking the derivative of the restricted cubic splines, the function changes and becomes negative after 10 years of age when describing the relationship between FEV_1 and age, while it remains always positive for $LCI_{2.5}$.

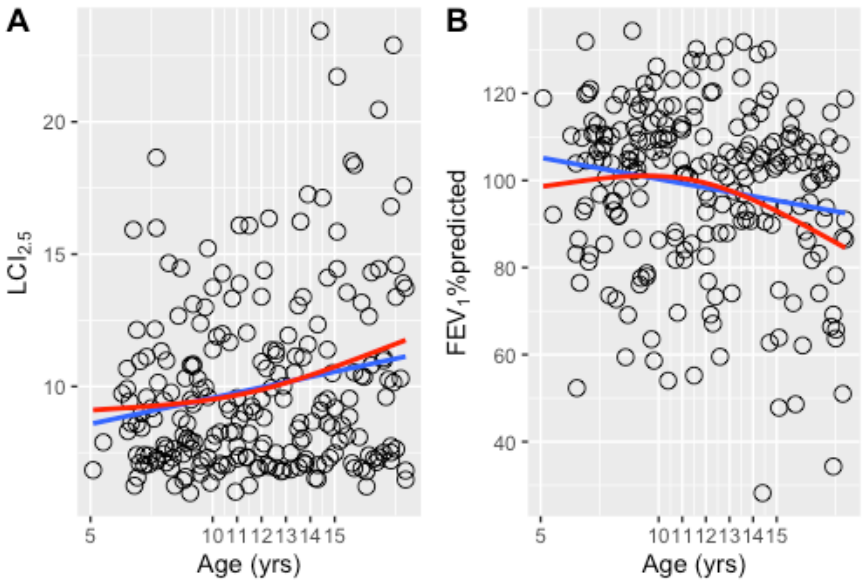

Figure 4. Red lines represent restricted cubic spline smoothing used to describe the relationship between A) $LCI_{2.5}$with age and between B) FEV_1% predicted and age. Restricted cubic splines used 3 knots, based on AIC. Blue lines represent linear regression line.

Missingness and imputation plausibility

Generally, nearly 86% of the sample present complete information in the data regarding spirometry and *Pseudomonas* infection, collected during outpatient evaluation of paediatric subjects with at least one $MBWN_2$ test. Spirometry-derived indices are missing in 20 subjects (≈8%), while data about microbial status are missing in 17 individuals (7%). In 3 individuals, both these variables were missing. Once these variables were plotted against age, a potential determinant in the mechanism of missingness, no observable relation was detected.

Figure 5 shows the distribution of FEV_1 % predicted and PSA as individual points and the distributions obtained after eight multiple imputations with 70 iterations for each imputation. Distributions of original and imputed variables look plausible, from a quantitative and clinical point of view, and do not show worrisome behaviours, as shown as well by FEV_1 %predicted density of the different imputed datasets used (Figure 6).

As far as the imputation of PSA missing data is concerned, chronic infection by *Pseudomonas aeruginosa* varies between imputed dataset, ranging from a minimum prevalence of 19.6% to maximum 21.6%.

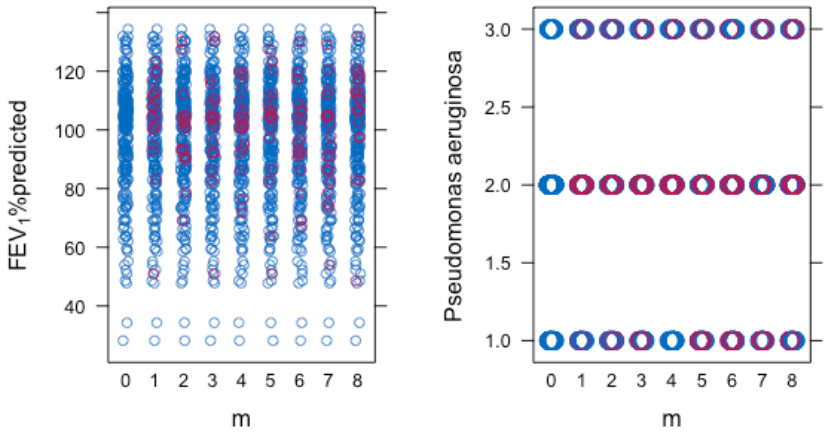

Figure 5. Scatterplot of FEV_1 (left) and Pseudomonas aeruginosa (right) against 8 imputations (1 to 8; 0 is the source dataset). Blue are observed and magenta are imputed data.

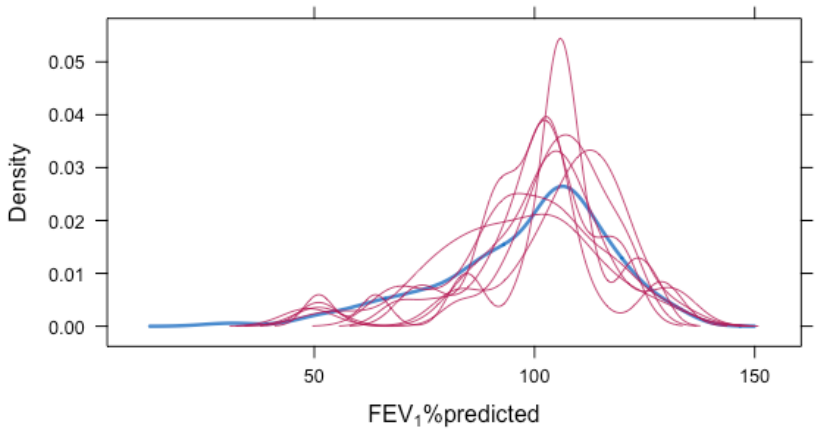

Figure 6. Kernel density estimates for the marginal distributions of the observed data (blue) and the m = 5 densities per variable calculated from the imputed data (thin red lines)

Association between lung function measures and selected variables

$LCI_{2.5}$ and FEV_1 %predicted are distributed as shown in Figure 7. Signs of departures from a normal distribution are evident, as also assessed by visual inspection of Q-Q plots. Shapiro-Francia test statistics denote these variables as not normally distributed (P<0.0001).

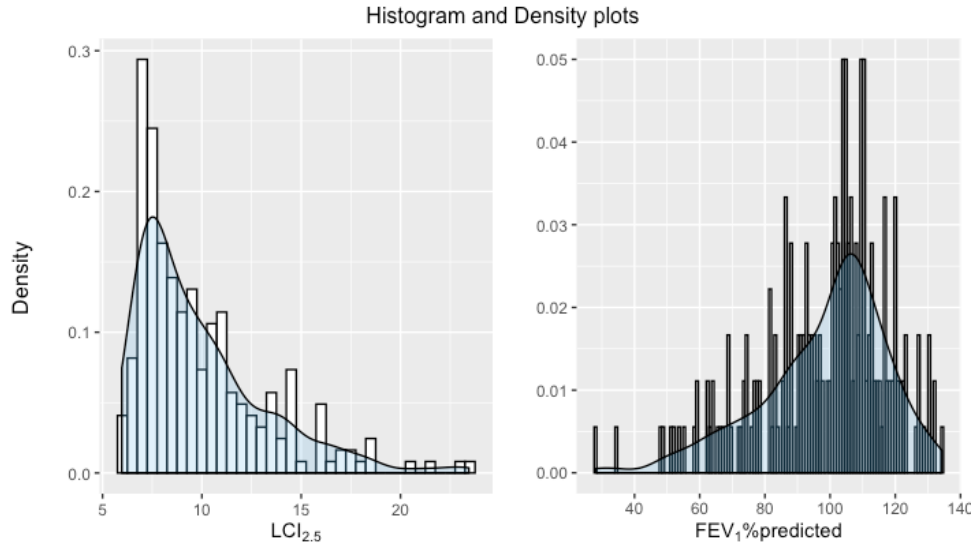

Figure 7. Graphs illustrate histograms of $LCI_{2.5}$ (left) and FEV_1 %predicted (right). Shadowed curves represent kernel density curves

Two distinct OLS multiple regression models were fitted using lung function as dependent variables and using age, sex, pancreatic status, *CFTR* genotype, *Pseudomonas aeruginosa* infection and BMI Z-score as covariates. Presence of diabetes was omitted as a covariate from the two models given that only four children presented this comorbidity and considering also well-known its impact on the course of the disease. Lung function, namely FEV_1 % predicted and $LCI_{2.5}$, was poorer in children with diabetes (n=4), 13.41 (4.9) and 87.7 (15.5)% respectively, compared to children without diabetes, 9.84 (3.25) and 98.6 (19.3)%. Figure 8 shows the distribution of the dependent variables in patients with and without diabetes.

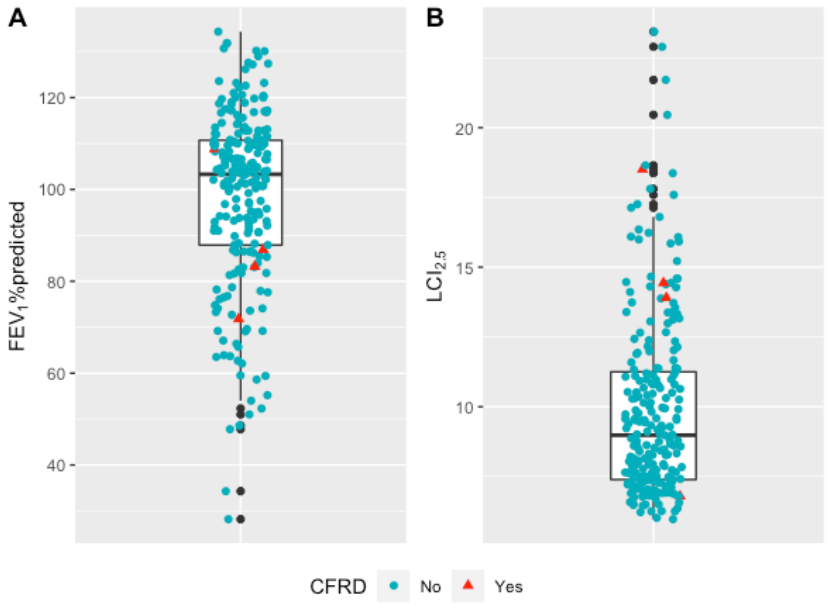

Figure 8. Box-plots of FEV_1 %predicted (A) and $LCI_{2.5}$ (B). Black dots define outliers.

Considering age as linear, OLS models were built as follows:

1) $$FEV_1\%predicted = b_0 + b_1(age) + b_2(sex) + b_3(mutation) + b_4(pancreatic\ function) + b_5(PS\ infection) + b_6(bmi)$$

2) $$LCI_{2.5} = b_0 + b_1(age) + b_2(sex) + b_3(mutation) + b_4(pancreatic\ function) + b_5(PS\ infection) + b_6(bmi)$$

Residuals analysis revealed a non-constant variance (Figure 9) and lack of normality, assessed by normal plot of residuals, particularly with the distribution of $LCI_{2.5}$. Moreover, Breusch–Pagan test for heteroskedasticity was significant only for $LCI_{2.5}$ (P=0.005) and not for FEV_1 %predicted (P=0.3641).

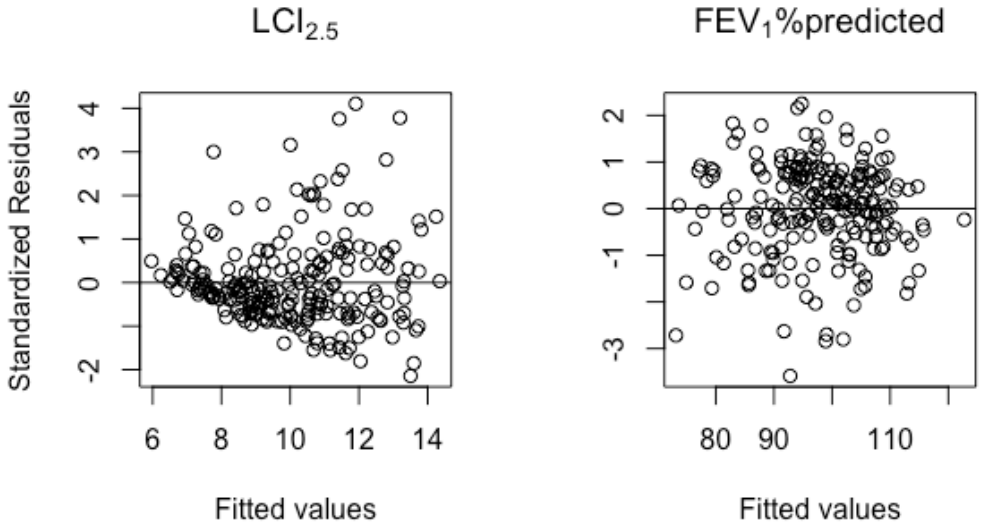

Figure 9. Graphs illustrate standardised residual plots of $LCI_{2.5}$ (left) and FEV_1 %predicted (right) versus fitted values.

Taking into consideration the relationship between the dependent variables and age, several attempts were made trying to summarize this relationship as non-linear in OLS models, introducing age-restricted cubic splines with 3 up to 6 knots. Table 4 reports the specific age modelling under the following regression equations:

3) $$FEV_1\%predicted = b_0 + spline(age, df) + b_2(sex) + b_3(genotype) + b_4(pancreatic\ function) + b_5(PS\ infection) + b_6(bmi)$$

4) $$LCI_{2.5} = b_0 + spline(age, df) + b_2(sex) + b_3(genotype) + b_4(pancreatic\ function) + b_5(diabetes) + b_5(PS\ infection) + b_6(bmi)$$

Table 4. Model selection statistics (AIC)

		FEV_1 %predicted		*$LCI_{2.5}$*	
	df	obs	AIC	obs	AIC
Age	10	*211*	1805.197	228	1124.466
Age-spline, 3 knots	11	*211*	1800.684	228	1124.600
Age-spline, 4 knots	12	*211*	1802.008	228	1126.449
Age-spline, 5 knots	13	*211*	1802.530	228	1128.163
Age-spline, 6 knots	14	*211*	1804.546	228	1130.114

df= degree of freedom; obs= number of observation; AIC = Akaike's information criterion

Based on the lowest AIC obtained, age modelled with a spline with 3 knots turned to be the preferred regression equations for FEV_1 %predicted while $LCI_{2.5}$ with age as linear had the lowest AIC.

The optimal Box-Cox estimated parameter for transforming the % predicted value of FEV_1 was 2, while -1.31 was the best Box-Cox estimated parameter to transform $LCI_{2.5}$. For subsequent analysis, FEV_1 %pred and $LCI_{2.5}$ were transformed to the square ($FEV_1\%predicted^{\,2}$) and to the inverse ($\frac{1}{LCI_{2.5}}$) transform, respectively. These transformations were deemed necessary in order to meet assumptions behind linear model.

Indeed, after transforming these dependent variables and fitting the models 3) and 4) again, the scatter of residuals improved, showing constant variance (Figure 10), and approximately normal distribution of residuals as well. Breusch–Pagan test for heteroskedasticity was no longer significant for $\frac{1}{LCI_{2.5}}$ (P=0.7495); residuals of FEV_1 %predicted and $LCI_{2.5}$ met a normal distribution under the Shapiro-Francia test statistics, P=0.3896 and P=0.2061 respectively.

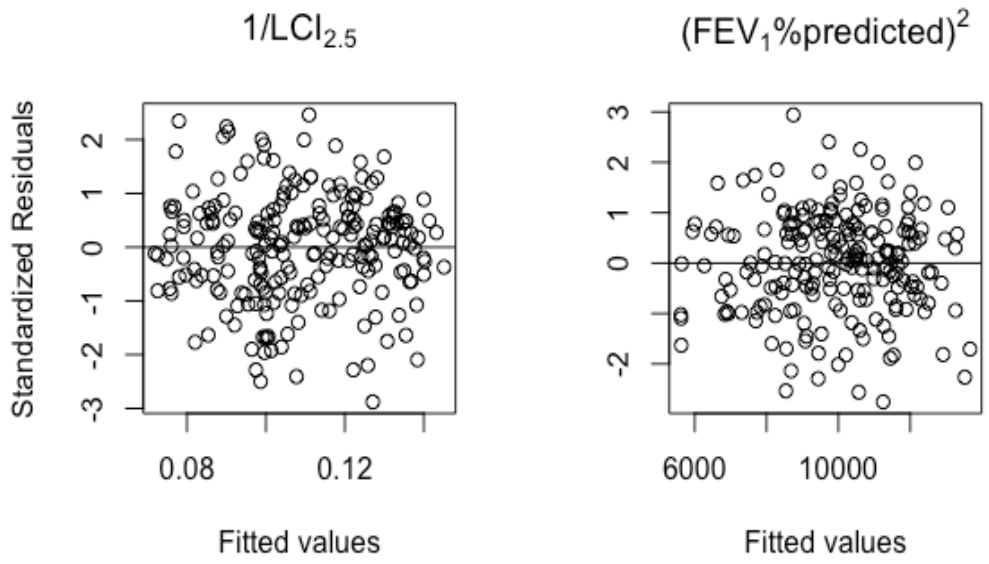

Figure 10. Graphs illustrate residual plots of transformed $LCI_{2.5}$ (left) and FEV_1 %predicted (right) versus predicted values.

Afterwards, the models 3) and 4) were run after multiple imputations (MI). The model with FEV_1 expressed as a square transform resulted in an increase in total variance due to missing information. For example, the relative increase in variance (RIV) for free *Pseudomonas aeruginosa* was 0.085, that is 8.5% larger than its sampling variance would have been had the data been complete. Particularly, estimated FMI for free and intermittent *Pseudomonas aeruginosa* colonization were 8.8% and 9.6%, respectively. Precision for all the estimates was above 98.7%.

As regards the model with $\frac{1}{LCI_{2.5}}$ as dependent variable, the RIV for the same microbial categories were 15.9% and 9.9% due to missingness. Precision for all the estimates was above 98.2%.

Residuals versus transformed dependent variables after MI are reported in Figure 11.

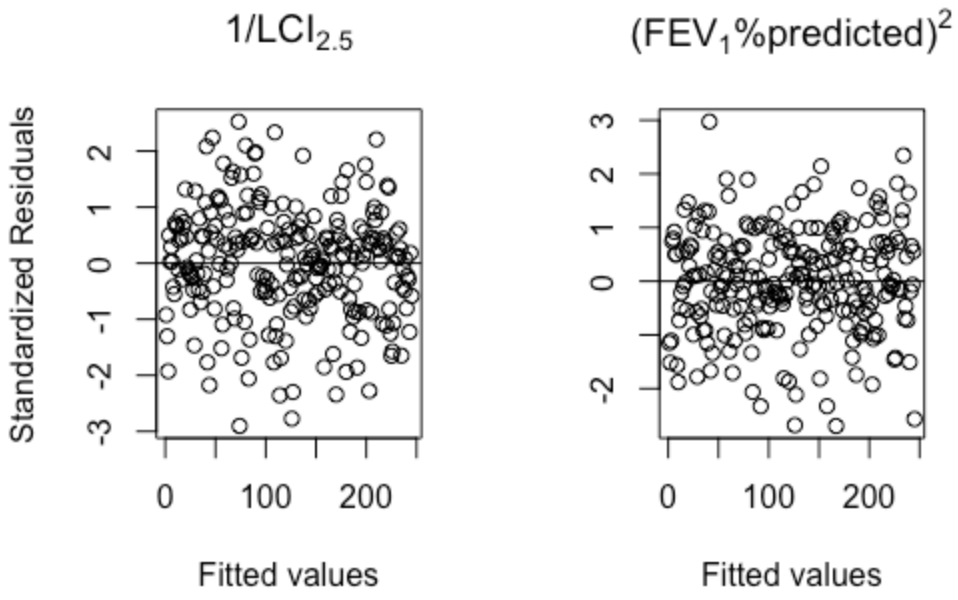

Figure 11. Graphs illustrate residual plots of transform LCI (left) and FEV_1 %predicted (right) versus predicted values, averaged over eight multiple imputations.

Generally, the imputed model with FEV_1 %predicted expressed as square function explains 22.8% of the variance of the dependent variable (Table 5). Concerning the covariates, there is no evidence of association between *CFTR* genotype and FEV_1 (P=0.084), whereas there is a significant association between *Pseudomonas aeruginosa* infection and airflow obstruction in the considered sample (P<0.001). This association is significant for patients with chronic infection compared to those free (b = -2099.12, P<0.001) or intermittent (b = -2004.43, P<0.001) (Figure 12).

Table 5. Association of prognostic factors with FEV_1: model-based coefficients standard errors in not imputed (top) and imputed dataset (bottom)

n=211	*b-Coefficients*	*Standard Error*	*P value*	*Pr(>F)*
Age-spline, 3 knots				0.003
Male vs Female	-70.44	427.78	0.869	
CFTR Genotype				0.076
F508del/other vs F508del/F508del	-1116.67	587.47	0.059	
Other/other vs F508del/F508del	-1435.35	646.16	0.027	
Pancreas Insufficiency	-1777.40	505.35	<0.001	
Pseudomonas aeruginosa infection				<0.001
intermittent vs chronic	1847.27	685.45	<0.001	
free vs *chronic*	2136.94	557.40	<0.001	
BMI, Z-score	841.16	242.43	<0.001	
Intercept	9084.38	1719.69		
n=245	*b-Coefficients*	*Standard Error*	*P value*	*Pr(>F)*
Age-spline, 3 knots				0.005
Male vs Female	97.98	412.85	0.81	
CFTR Genotype				0.084
F508del/other vs F508del/F508del	-1039.66	580.96	0.246	
Other/other vs F508del/F508del	-1433.18	634.65	0.563	
Pancreas Insufficiency	-1987.32	483.56	<0.001	
Pseudomonas aeruginosa infection				<0.001
intermittent vs chronic	2004.43	674.58	<0.001	
free vs *chronic*	2099.12	550.69	<0.001	
BMI, Z-score	800.43	229.98	<0.001	
Intercept	8942.30	1677.84		

As well pancreatic function has a negative association with the response variable, while a better nutritional status (BMI) is positively associated with lung function (b = 800.43, P<0.001). The association of age with FEV_1 %predicted is slightly increasing until 12, decreasing afterwards, as described in Figure 12, with children chronically infected by *Pseudomonas aeruginosa* showing the worst trend.

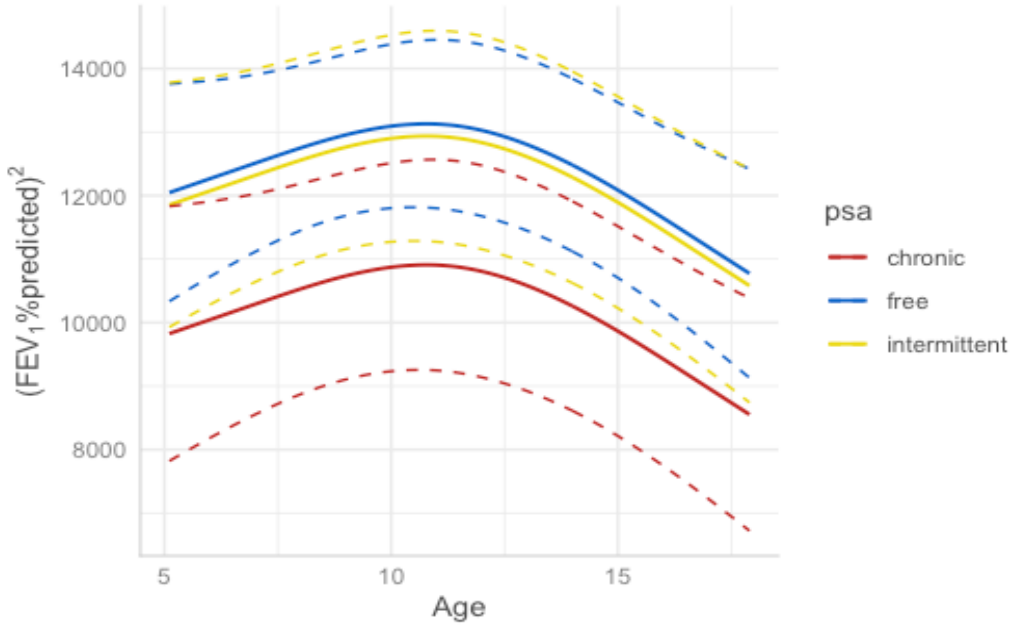

Figure 12. Solid lines represent fitted values obtained from the model with transform FEV_1% predicted after multiple imputation on each dataset, according to PSA status; dot lines represent 95% confidence intervals of model estimates. Here it is represented the marginal model estimated on imputed dataset #2, for a female subject, carrying two *F508del* variants and with BMI equals to -0.5.

The other model with $\frac{1}{LCI_{2.5}}$ as dependent variable explains a larger amount of variance compared to the previous one (37.6% of the total variation of $\frac{1}{LCI_{2.5}}$). As reported by the estimated coefficients in table 6, this model as well revealed a significant association between pancreatic status and the dependent variable.

Table 6. Association of prognostic factors with $LCI_{2.5}$: model-based coefficients standard errors in not imputed (top) and imputed dataset (bottom)

n=228	*b-Coefficients*	*Standard Error*	*P value*	*Pr(>F)*
Age	-0.002	0.000	<0.001	
Male vs Female	-0.001	0.003	0.685	
CFTR Genotype				0.004
F508del/other vs F508del/F508del	0.005	0.004	0.236	
Other/other vs F508del/F508del	0.014	0.005	0.002	
Pancreas Insufficiency	-0.019	0.004	<0.001	
Pseudomonas aeruginosa infection				<0.001
Intermittent vs chronic	0.009	0.005	0.072	
free vs chronic	0.020	0.004	<0.001	
BMI,Z-score	0.004	0.002	0.011	
Intercept	0.121	0.008		
n=245	*b-Coefficients*	*Standard Error*	*P value*	*Pr(>F)*
Age	0.195	0.000	0.003	
Male vs Female	-0.001	0.000	0.801	
CFTR Genotype				0.013
F508del/other vs F508del/F508del	0.006	0.004	0.170	
Other/other vs F508del/F508del	0.013	0.005	0.005	
Pancreas Insufficiency	-0.022	0.004	<0.001	
Pseudomonas aeruginosa infection				<0.001
Intermittent vs chronic	0.010	0.005	0.056	
free vs chronic	0.019	0.004	<0.001	
BMI,Z-score	0.004	0.002	0.015	
Intercept	0.121	0.008		

A positive association between *Pseudomonas aeruginosa* and the lung clearance index is also detectable under this model, being chronically infected worse than having an intermittent colonisation (Figure 13). The model did not account for a difference in mean $\frac{1}{LCI}$ among those children with an intermittent *versus* chronic colonization. The association of age with transform $LCI_{2.5}$ was significant (P=0.003), as it was the association between genotype.

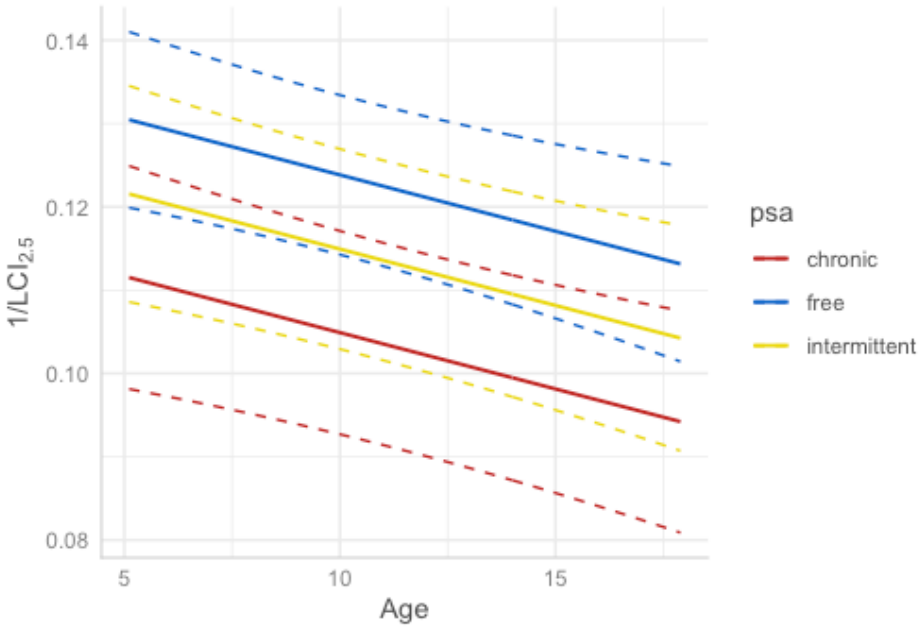

Figure 13. Solid lines represent fitted values obtained from the model with transform $LCI_{2.5}$ after multiple imputation on each dataset, according to PSA status; dot lines represent 95% confidence intervals of model estimates. Here it is represented the marginal model estimated on imputed dataset #2,, for a female subject, carrying two *F508del* variants and with BMI equals to -0.5.

Profiling paediatric patients according to their clinical characteristics

Characteristics of the 125 paediatric subjects evaluated at least twice during outpatient follow-up are shown in Table 7.

Table 7. Sample characteristics

Visits	*I*	*II*	*III*	*IV*
Subjects, *nr*	125	125	47	15
Age, *yrs*	11.4 (2.8)	12.9 (2.8)	14.3 (2.1)	14.7 (1.8)
BMI, Z-score	-0.6 (0.9)	-0.5 (0.9)	-0.6 (1.0)	-0.9 (0.7)
Male/Female, *nr*	64 / 61	64 / 61	23 / 24	11 / 4
F508del/*F508del*, *nr(%)*	31 (24.8)	31 (24.8)	12 (25.5)	6 (40)
F508del/ other, *nr(%)*	50 (40.0)	50 (40.0)	18 (38.3)	6 (40)
other/other, *nr(%)*	44 (35.2)	44 (35.2)	17 (36.2)	3 (20)
CFRD, *nr(%)*	1 (0.8)	1 (0.8)	1 (2.1)	-
Pancreatic Insufficiency, *nr(%)*	74 (59.2)	74 (59.2)	30 (63.8)	10 (66.7)
Pseudomonas aeurginosa chronic infection*, *nr(%)*	25 (20)	18 (14.4)	8 (17.0)	4 (26.7)

*Values are expressed as absolute number(percentage); * 8 microbiological data are missing*

Children show good respiratory conditions, as these were evaluated during the first outpatient visit (Table 8). Despite a little decrease in airflow obstruction, as shown by mean Z-score of FEV_1, mean FEV_1 %predicted has remained above 90% throughout test occasions. During the first evaluation, no children expressed severe lung disease (FEV_1% < 40% predicted), nevertheless 15 children (12.7%) were considered below their lower limit of normal (LLN). The percentage of children with FEV_1 below LLN did not vary a lot through follow-up: 13 (11.7%) at second evaluation, 6 (13.3%) at third and 4 (26.7%) at fourth evaluation. Two patients evaluated four times were all below their LLN.

Table 8. Pulmonary characteristics

Visits	*I*	*II*	*III*	*IV*
Subjects, *nr*	125	125	47	15
FEV_1, *% predicted*	100.4 (17.4)	97.8 (16.7)	97.4 (20.1)	93.5 (17.8)
FEV_1, Z-*score*	0.1 (1.5)	-0.2 (1.4)	-0.2 (1.7)	-0.5 (1.5)
FVC, *% predicted*	107.4 (14.7)	105.5 (14.2)	103.3 (16.5)	104.1 (13.5)
FVC, Z-*score*	0.6 (1.2)	0.4 (1.2)	0.3 (1.4)	0.3 (1.1)
$LCI_{2.5}$	10.07 (2.98)	9.92 (3.32)	9.50 (2.75)	10.05 (2.06)
$LCI_{2.5}$, *CV%*	4.4 (2.4)	4.0 (2.3)	4.4 (2.2)	3.8 (2.7)
$LCI_{2.5}$, *Z-score*[¶]	5.22 (1.78 – 6.67)	4.16 (1.47 – 9.98)	3.62 (1.07 – 8.31)	5.49 (2.9 – 9.39)
Scond[*VT]	0.066 (0.030)	0.141 (0.799)	0.061 (0.027)	0.066 (0.019)
Scond, *CV%*	26.9 (25.2)	27.9 (25.8)	26.5 (24.3)	21.2 (15.8)
Sacin[*VT]	0.138 (0.109)	0.134 (0.104)	0.123 (0.096)	0.139 (0.117)
Sacin, *CV%*	31.9 (26.5)	34.1 (27.9)	30.8 (24.3)	35.3 (29.9)

Spirometry data belong to 118 patients at first evaluation, 111 to second follow-up and to 45 patients at third evaluation. Sacin was not detected on 1 and 5 patients during first, third and second evaluation, respectively. All displayed values are expressed as mean(standard deviation). [¶]median and interquartile range. One patient was evaluated five times (not reported).

Average $LCI_{2.5}$ did not show large changes over time, considering that patients were followed-up after a period of time comprised between 469 and 636 days, first and third quartiles respectively.

If compared to those children who performed at least one $MBWN_2$ (Table 1 and Table 2, pg. 10 and 11), no substantial clinical difference can be reported.

Agglomerative nesting was the adopted clustering algorithm used to generate CF phenotypes. Table 9 displays the several indices taken into consideration. First, the number of clusters was 2 according the majority of indices evaluated. Despite the best performance among all the indices shown by *hclust*, the analysis of dendrogram revealed a severe imbalance in the number of subjects within each cluster, 2 patients in the first and 123 in the second cluster, therefore this approach was discarded as clinically useless. AGNES showed a slightly better performance in the validity of its clustering structure against poorer stability, compared to PAM.

Table 9. Validity and stability indices and other descriptive clusters properties

	AGNES (k=2)	PAM (k=2)	Hclust (k=2)	*Reference*
Linkage	Ward	-	Average	
Metric	Gower	Gower	Gower	
Correlation coefficient	0.59	-	0.60	*>0.80*
Agglomerative coefficient	0.96	-	-	
Silhouette Index	0.33	0.31	0.35	*>0.50*
Dunn Index	0.20	0.10	0.27	*Highest*
Average within	0.22	0.25	0.29	*Lowest*
Average between	0.35	0.35	0.44	*Largest*
APN	0.0961	0.0113	0.0961	*Lowest*
AD	15.115	12.415	15.115	*Lowest*
ADM	0.1501	0.0168	0.1501	*Lowest*
FOM	0.0987	0.0840	0.0987	*Lowest*

APN = Average proportion of non-overlap; AD = Average Distance; ADM = Average distance between means; FOM = Figure of merit.

The clustering hierarchy as a tree diagram obtained by applying AGNES algorithm on patients with at least two $MBWN_2$ tests, at their first evaluation, is shown in figure 14.

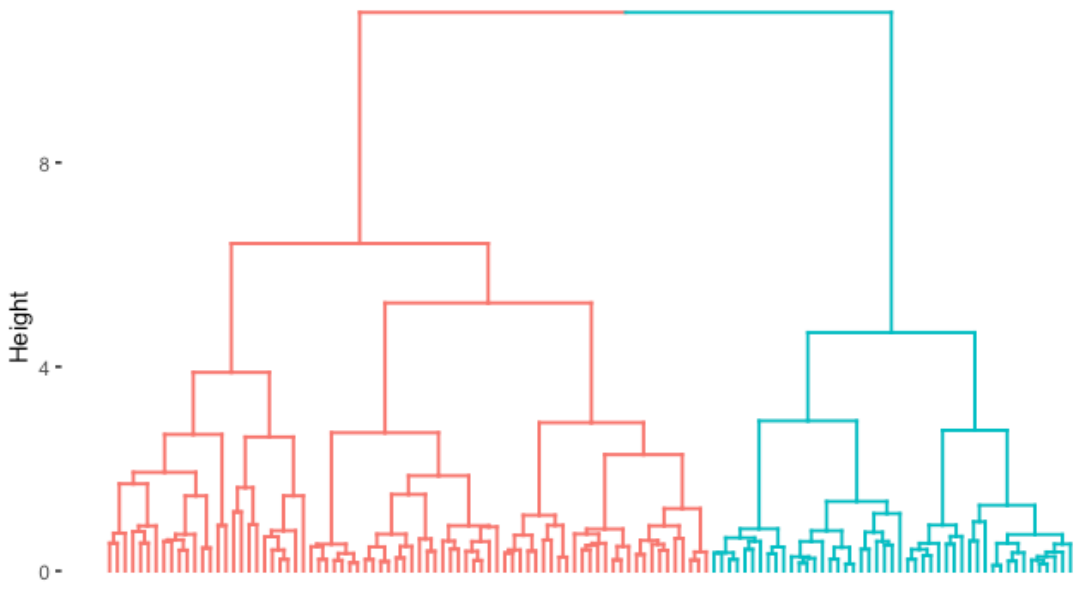

Figure 14. Dendrogram. Vertical axis represents the dissimilarity between clusters. The horizontal axis represents patients

The number of identified clusters was also considered reasonable after looking at the dendrogram.

Table 10 describes the clinical characteristics for subjects included in each cluster.

Table 10. Sample characteristics stratified by clusters

		Cluster 1	Cluster 2	p-value
Subjects, *nr*		78	47	
Age, *yrs*		11.09 (2.97)	11.80 (2.31)	0.156
Age range, *yrs*		5.6 - 16.8	7.7- 16.3	
Sex, female		36 (46.2)	25 (53.2)	0.466
BMI, *Z-score*		-0.68 (0.78)	-0.37 (1.00)	0.073
Pancreatic Insufficiency		74 (94.9)	-	<0.001
CFRD		1 (1.3)	-	1.000
CFTR Genotype				<0.001
	F508del/other	32 (41.0)	18 (38.3)	
	F508del/*F508del*	30 (38.5)	1 (2.1)	
	Other/other	16 (20.5)	28 (59.6)	
Pseudomonas aeurginosa				<0.001
	chronic	25 (32.9)	0 (0.0)	
	free	41 (53.9)	35 (85.4)	
	intermittent	10 (13.2)	6 (14.6)	

FEV_1, *% predicted*		95.26 (17.96)	109.01 (12.41)	<0.001
FEV_1, Z-*score*		-0.38 (1.52)	0.79 (1.08)	<0.001
FVC, *% predicted*		104.39 (15.98)	112.36 (10.68)	0.002
FVC, Z-*score*		0.35 (1.34)	1.03 (0.90)	0.001
$LCI_{2.5}$		11.33 (2.91)	7.97 (1.57)	<0.001
$LCI_{2.5}$, *Z-score,* median (IQR)		8.32 (8.44)	0.98 (2.98)	<0.001
Sacin^{*VT}		0.16 (0.12)	0.10 (0.08)	0.001
Scond^{*VT}		0.08 (0.02)	0.05 (0.03)	<0.001
Pulmonary Exacerbations*, median (IQR)		3 (1-4)	0 (0-1)	<0.001
Hospitalization*				<0.001
	1	15 (19.2)	-	
	⩾2	4 (5.1)	-	

*Values are expressed as absolute number(percentage) or mean(sd), where not differently expressed. * Reference time period is twelve months preceding $MBWN_2$ test*

One third of the considered sample fell into cluster #2, made by young individuals of the same age of cluster #1, but less severe. Indeed, cluster #1 is characterized by the co-presence of several negative known prognostic factors, statistically different in their distribution from cluster #2. For example, *F508del* variant is almost absent in cluster #2 and none presents pancreatic insufficiency. Moreover, individuals in cluster #2 are classified as free from *Pseudomonas aeruginosa* colonization whereas all youngsters with chronic infection by PSA fell into cluster #1.

The sole analysis of air flow obstruction in terms of FEV_1 %predicted returns all patients in the two clusters in the range of moderate to mild/normal air flow obstruction, with a median (IQR) FEV_1 equals to 96.4 (85.7 – 109.4) %predicted in the cluster #1 and 106.2 (103.2 – 119.3) % predicted in cluster #2. Particularly, 90.5% and 100% of patients, respectively in clusters #1 and #2, could be considered having mild/normal FEV_1; again, these percentages vary if one adopts Z-score of FEV_1, under which case the percentage of patients with a normal lung function drops to 81.1% and to 97.7% in the first and second cluster, respectively. Moreover, by adopting Z-score of FEV_1, the two clusters

show exactly an opposite behaviour, with a median (IQR) FEV_1 equals to -0.30 (-1.19 – 0.83) Z-score in the first cluster and 0.53 (0.28 – 1.64) in the second cluster.

The analysis of $LCI_{2.5}$ improves the understanding of lung status and clearly classifies as abnormal, i.e. $LCI_{2.5}$ >ULN, the majority of patients in cluster #1, 92.3% compared to 42.6% in cluster #2. At a population level, 95%CI for the mean $LCI_{2.5}$ in cluster #1 goes from 10.7 to 12.0, whereas 95%CI for the mean $LCI_{2.5}$ in cluster #2 is much lower, namely 7.5 to 8.4. Acinar and conductive ventilation appear to be more homogenous in cluster #2 compared to cluster #1. Generally, the two clusters differentiated one from another by a mean $LCI_{2.5}$ difference of 3.37 (95%CI: 2.57 to 4.16) and by a mean FEV_1 difference of 1.2 (95%CI:0.7 to 1.6) Z-score. Under these clusters, no sex difference is detectable (P=0.466) nor any significant difference in nutritional status, -0.34 (95%CI: -0.65 to 0.03) BMI Z-score. Age span is pretty much identical between the two partitions.

Longitudinal association between lung function measures and selected variables between phenotypes

Considering that only 14 patients in cluster #1 and one patient in cluster #2 performed 4 visits or more during the considered period of time, the analysis of $LCI_{2.5}$ variation over time covers a 3 visits-span. On average, children were evaluated for a second time after 1.6 years during a regular outpatient visit; for 47 children, the third evaluation occurred after 2.8 years from the first evaluation and 1.4 years from the second evaluation.

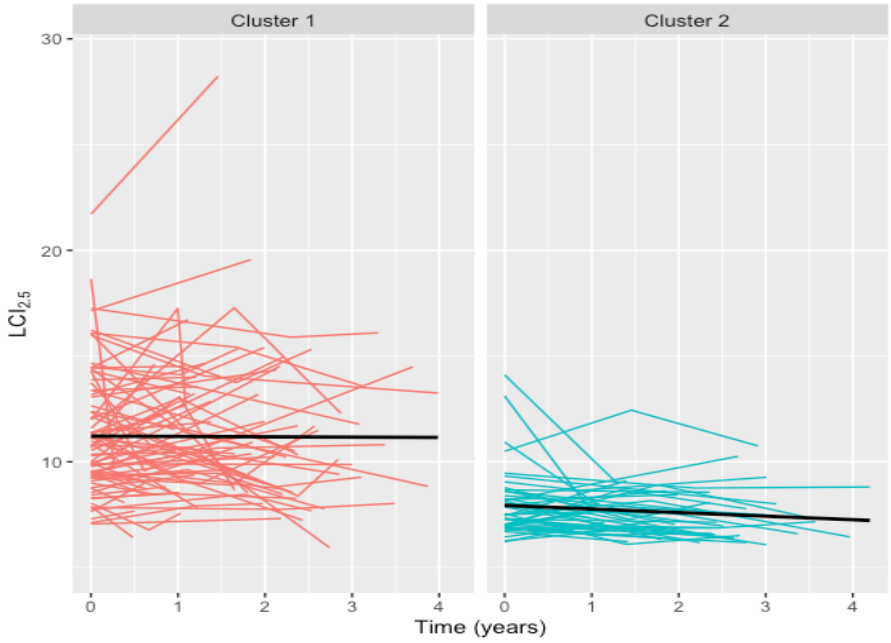

Figure 15: $LCI_{2.5}$ profile (spaghetti plot) for children over years. Solid line describes the linear regression used to visually describe the relationship between time and $LCI_{2.5}$

From figure 15, three things stand out. The first is about one patient in cluster #1, whose baseline $LCI_{2.5}$ is above 20 and further increased at the next follow-up. This represents an unusual observation belonging to a young girl aged 15 years, whose FEV_1 deteriorated in one year, going from 47.8% to 32.3 %predicted. Considering the poor meaning of such a high measure of $LCI_{2.5}$, we decided to remove this patient from next analyses.

Secondly, patients in cluster #1 have a lot of variability in their ventilation inhomogeneity basal measurement and, thirdly, children in cluster #2 seem more stable over time.

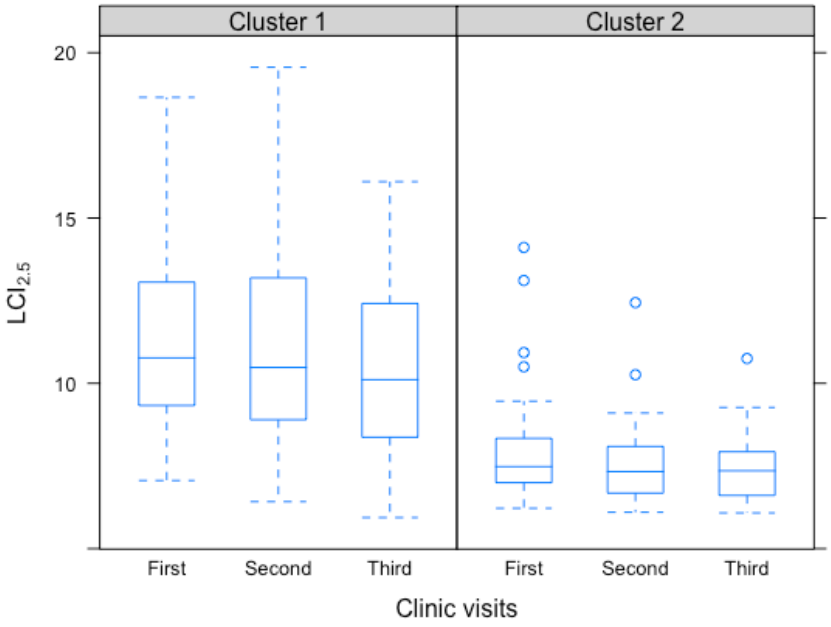

Figure 16: Box-and-whiskers plots for $LCI_{2.5}$ by cluster and time.

Indeed, figure 16 shows the patterns implied by the sample median. The variability is larger in cluster #1 compared to cluster #2, and it seems almost constant over time.

The mixed-effect model was built as follows, allowing an interaction between visits and the two identified phenotypes (i.e. clusters):

$$5)\ LCI_{2.5\,ij} = b_0 + b_1(visit_{ij}) + b_2(cluster_i) + b_{12}(visit_{ij}\ X\ cluster_i) + U_{ij}$$

where U_{ij} is the patient-specific random effect. Based on the LR test ($P<0.001$), the final model considering a heterogeneous residual variance structure within cluster indicates a better fitting of the model. This choice is in agreement with the computed value of 1237.98 of AIC, which is slightly smaller than 1263.15 obtained from the model with constant variance.

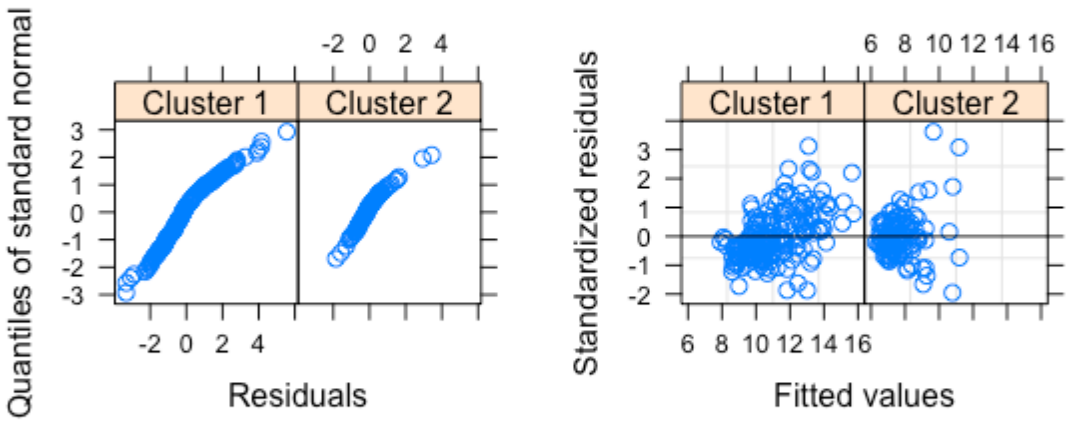

Figure 17. Graphs illustrate normal Q-Q of the conditional Pearson residual for each cluster (left) and scatterplots (right) of the conditional Pearson residual versus fitted values from equation 5.

Figure 17 displays the diagnostic plots used to assess the goodness of fit of model 5. The left plot shows little deviation from linear trend whereas plot on the right shows that residual variance is slightly different for higher values, especially for observations contained in cluster #2. It is worth considering that diagnostic plot from this model with heterogeneous variance is qualitatively better and revealed a smaller number of outliers (n=7), i.e. residuals larger than the 97.5th percentile of standard normal distribution, compared to the model with constant variance (n=25).

Considering that outliers are present in both clusters and at first and second timepoints (figure not shown), and that dropping them from the final model did not change the overall magnitude and interpretation of estimates, these were not discarded. The resulting $LCI_{2.5}$ profile across time is displayed in Figure 18, that identify outliers and the negligible slope changing.

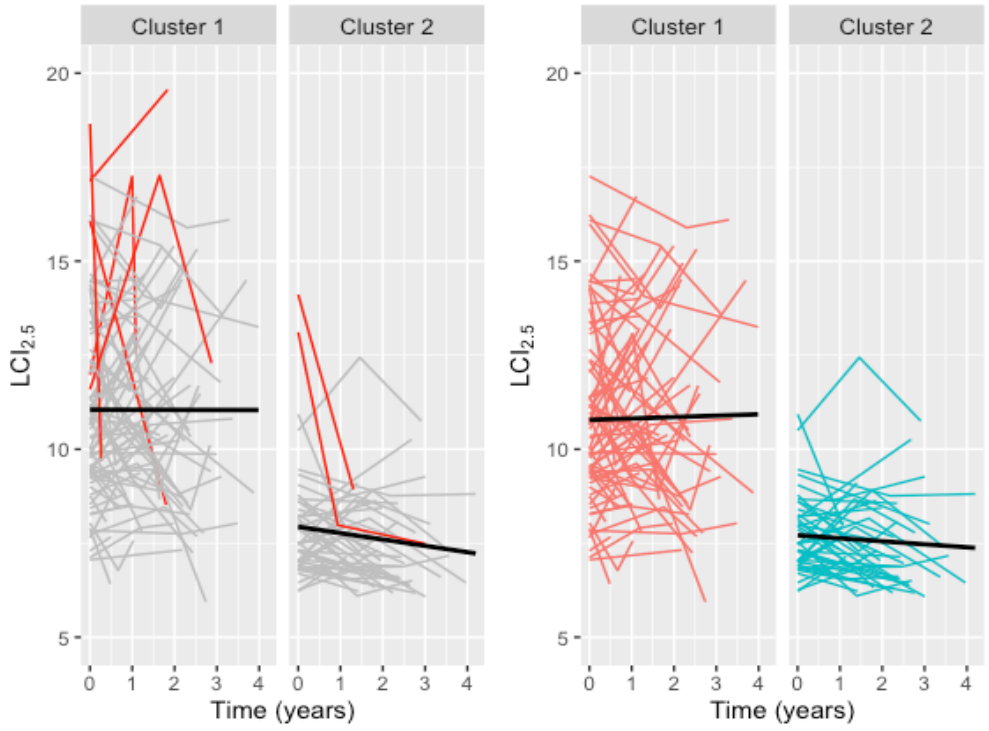

Figure 18. $LCI_{2.5}$ profile (spaghetti plot) for children over years. Plot on the left highlights the 7 observations considered as outliers (red lines) whereas plot on the right displays the variation across visits if outliers were removed. Solid black line describes the linear regression used to visually describe the relationship between time and $LCI_{2.5}$

Also, in order to exclude an impact of non-Gaussian distribution of residuals on the precision of the estimates and its interpretation, a model with the inverse of $LCI_{2.5}$ was fitted, yielding to the same clinical conclusions. For the sake of interpretability, results are therefore presented with untransformed $LCI_{2.5}$.

Table 11. REML-based parameter estimates for $LCI_{2.5}$ with subject-specific random intercepts

	b-Coefficients (SE)	p-value	*Pr(>F)*
Fixed effects			
Intercept	11.20 (0.28)	<0.001	
Cluster			
Cluster 2 *versus* Cluster 1	-3.23 (0.39)	<0.001	
Visits			
Visit 2 *versus* Visit 1	-0.08 (0.29)	0.7805	0.1265
Visit 3 *versus* Visit 1	-0.80 (0.40)	0.0478	
Cluster X Visit			
Δ Visit 1 *versus* Δ Visit 2	-0.31 (0.35)	0.3720	0.5023
Δ Visit 1 *versus* Δ Visit 3	0.21 (0.50)	0.6841	

SE = standard error; Δ = difference betwen clusters

Results from the mixed-effect model with heterogenous residual variance are displayed in table 11. Generally, difference between children explain 46% of the variance left over after the variance explained by clustering and follow-up visits. There is evidence of association between clustering and ventilation inhomogeneity in this sample but no evidence of different effect of time between clusters (P=0.5023). $LCI_{2.5}$ is on average 3.23 (95%CI: 2.46 to 4) lower in cluster #2 than in cluster #1 at baseline. At the second follow-up visit, the absolute difference between clusters is 3.54 (95%CI 2.77 to 4.31) while at the third follow-up it is 3.02 (95%CI 1.98 to 4.08). These differences are statistically significant and clinically meaningful.

As regards follow-up visits, the model describes a non-significant variation over time (P=0.1265). In cluster #1, the slope of $LCI_{2.5}$ at the second follow-up changes by -0.08 (95%CI: -0.65 to 0.49) and by -0.8 (95%CI: -1.6 to -0.01) at the third visit, compared to the baseline. In cluster #2, the slope of $LCI_{2.5}$ at the second follow-up changes by -0.39 (95%CI: -0.78 to 0.01) and by -0.2 (95%CI: -0-8 to 0.39) at the third visit, compared to the baseline.

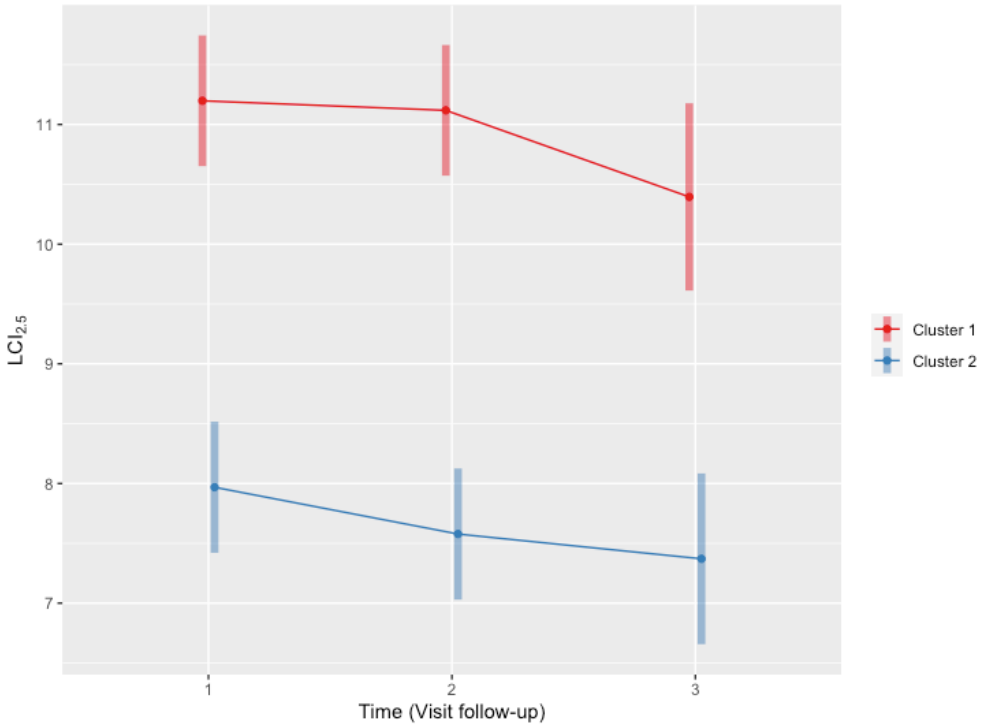

Figure 19. Solid lines represent fitted values obtained from the marginal model, with visits follow-up expressed as discrete time. Vertical bar lines represent 95% confidence intervals.

An adjusted model with colonization by *Pseudomonas aeruginosa* (free *versus* intermittent or chronic), age as linear, genotype (*F508del* homozygotes, heterozygotes and other variants), pancreatic status, sex, BMI Z-score was fitted to explore also the association of $LCI_{2.5}$ with clinical markers of CF disease. The modelling also accounted for the number of hospitalizations and pulmonary exacerbations experienced in the twelve months before the $MBWN_2$ test.

An association between *Pseudomonas aeruginosa* and the lung clearance index is also detectable under this adjusted model (P=0.0037), showing that being free from *Pseudomonas aeruginosa* lowers $LCI_{2.5}$ by 0.82 (95%CI: -1.36 to -0.27). Other evidences of association come with age (P=0.009), which increases $LCI_{2.5}$ by 0.17 for every one-year (95%CI: 0.04 to 0.3), with pancreatic status (P<0.001) and with BMI Z-score (P=0.0017). Particularly, those with pancreatic sufficiency have lower $LCI_{2.5}$ compared to those with pancreatic insufficiency (-1.43, 95%CI: -2.11 to -0.75), whereas

nutritional status (BMI Z-score) is negatively associated with ventilation inhomogeneity (b = -0.533, P=0.0017).

As regards the evaluation of FEV_1 %predicted, figure 20 summarise the trend across follow-up visits, showing a large amount of heterogeneity between children in both clusters, however smaller in cluster #2. From the visual inspection of the spaghetti plot, it seems that cluster #1 and #2 show a decreasing trajectory of FEV_1 %predicted over time, more pronounced in cluster #1.

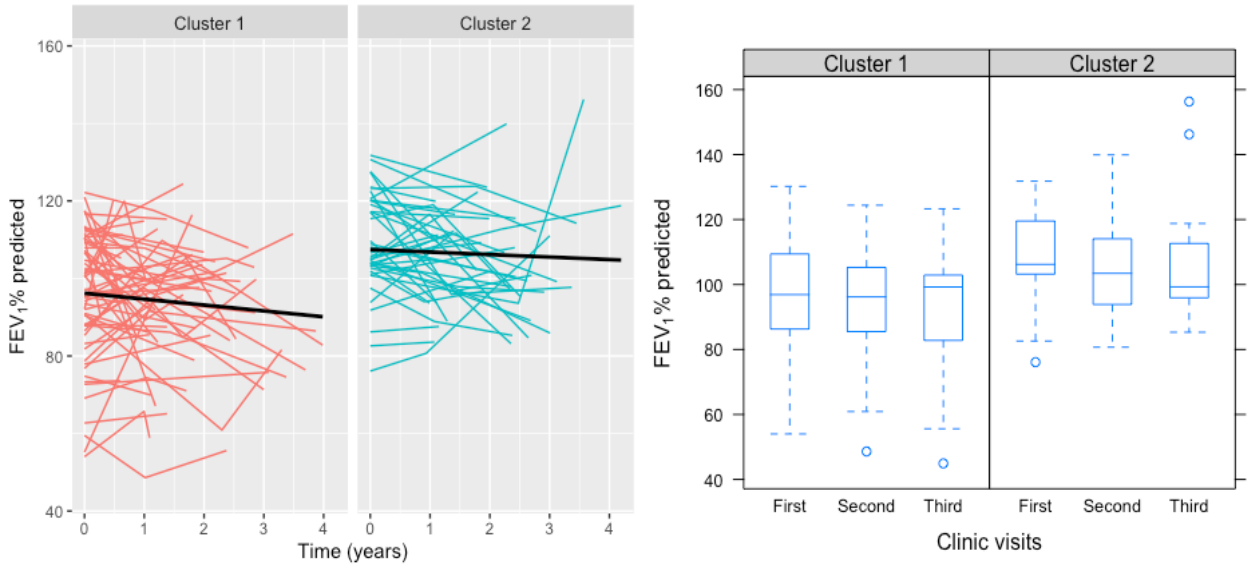

Figure 20: FEV_1 % predicted profile (spaghetti plot) for children across time. Solid line describes the linear regression used to visually describe the relationship between time and FEV_1 % predicted (left). On the right, Box-and-whiskers plots for FEV_1 % predicted by cluster and time.

The modelling of FEV_1% predicted was based on the following parametrization:

$$6)\ FEV_1\%\ predicted_{ij} = b_0 + b_1(visit_{ij}) + b_2(cluster_i) + b_{12}(visit_{ij}\ X\ cluster_i)\ + U_{ij}$$

where U_{ij} is the patient-specific random effect. Based on the LR test (P=0.5392), the final model considered a homoscedastic residual variance (AIC 2235.501), and therefore a general correlation structure was used. No worrisome departure from normality were detected in the diagnostic plots. Seven observations were identified as outliers and again were not excluded from the final model.

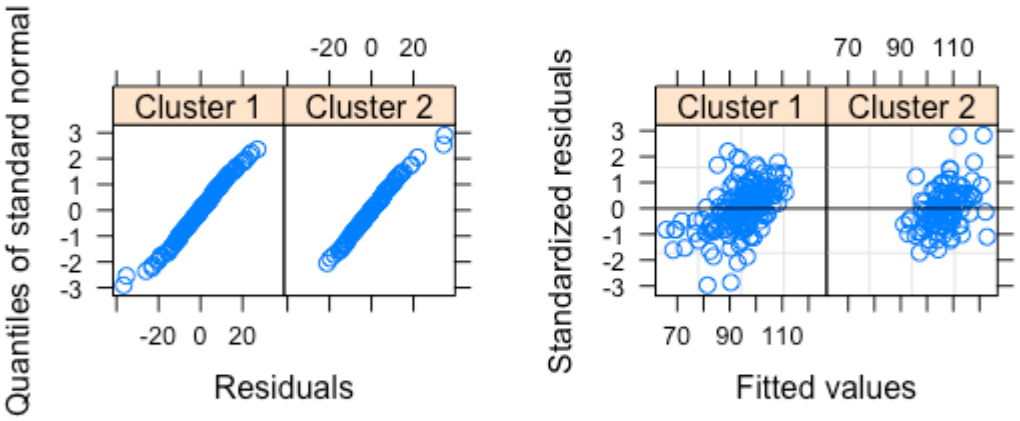

Figure 21. Graphs illustrate normal Q-Q of the conditional Pearson residual for each cluster (left) and scatterplots (right) of the conditional Pearson residual versus fitted values from equation 6.

Difference between patients explained a smaller portion of variance (i.e., 39.5%) compared to the model for $LCI_{2.5}$. Table 12 reports the coefficients for the marginal model of FEV_1 % predicted.

Table 12.REML-based parameter estimates for FEV_1 with subject-specific random intercepts

	b-Coefficients (SE)	p-value	*Pr(>F)*
Fixed effects			
Intercept	95.8 (1.8)	<0.001	
Cluster			
Cluster 2 *versus* Cluster 1	13.3 (3.0)	<0.001	
Visits			
Visit 2 *versus* Vitis 1	-1.1 (2.1)	0.5938	0.2415
Visit 3 *versus* Visit 1	-3.8 (2.9)	0.1843	
Cluster X Visit			
Δ Visit 1 *versus* Δ Visit 2	-3.2 (3.4)	0.3469	0.4408
Δ Visit 1 *versus* Δ Visit 3	2.5 (4.8)	0.6088	

SE = standard error; Δ = difference betwen clusters

Also, for FEV_1 %predicted we can observe evidence of a significative difference between clusters however without evidence of variation across visits (P=0.2415). FEV_1 is higher on average 13.3 % points predicted (95%CI: 7.3 to 19.2 %pred.) in cluster #2, at baseline. The overall estimated marginal means are displayed in figure 22.

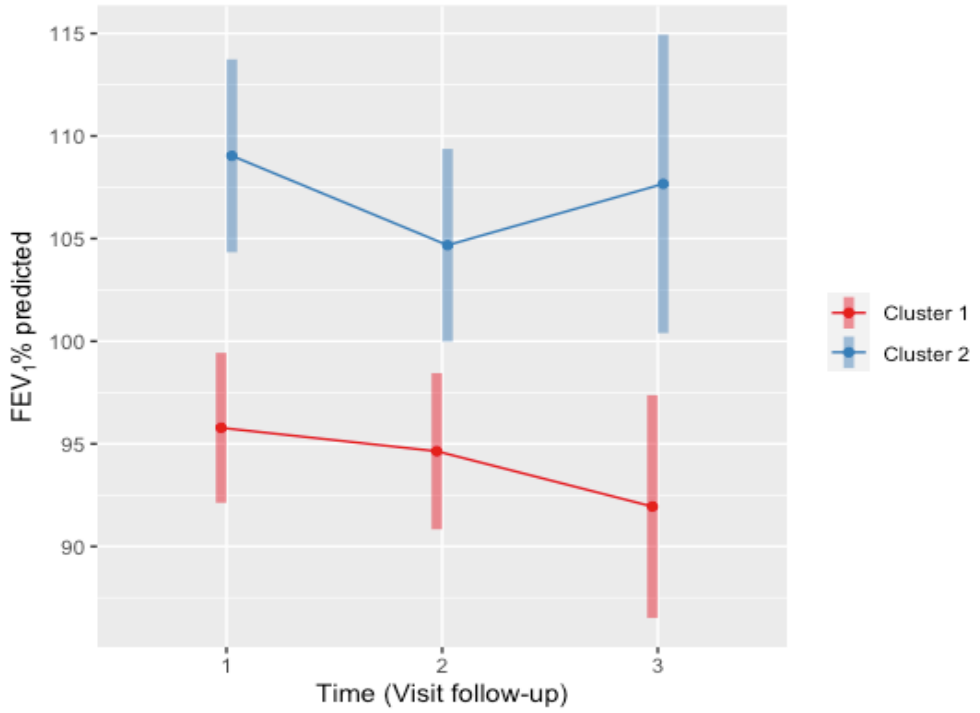

Figure 22. Solid lines represent fitted values obtained from the marginal model, with visits follow-up expressed as discrete time. Vertical bar lines represent 95% confidence intervals.

The same adjusted model used for $LCI_{2.5}$ was fitted to explore also the association of FEV_1 with clinical markers of CF disease. Based on AIC, age was expressed as non-linear using restricted cubic spline with 3 knots. A positive association between pancreatic function and FEV_1 was found (b=7.58, P=0.0048). In the same model, nutritional status (BMI Z-score) is positively associated with airflow obstruction (b = 4.77, P<0.001). No evidence of association was found between *Pseudomonas aeruginosa* or age and FEV_1 %predicted.

Children under therapy with *CFTR*-modulator agents

By the end of the present data collection, 11 children were under therapy with Orkambi® or Ivacaftor®. These children fell into cluster #1, and we were able to track longitudinal variation in their pulmonary functions only for three of them. Variation in their pulmonary function is summarised in Table 13.

Table 13. Pulmonary functions for children under therapy with modulators agents.

		Baseline visit	Follow-up	Relative Difference %
#1				
	FEV_1 % predicted	65.7	58.9	-10.4
	$LCI_{2.5}$	17.25	12.54	-27.3
	Age, *yrs*	11.4	11.5	
#2				
	FEV_1 % predicted	92.1	120.1	30.4
	$LCI_{2.5}$	7.89	6.42	-18.6
	Age, *yrs*	5.5	6.1	
#3				
	FEV_1 % predicted	119.9	99.2	-17.3
	$LCI_{2.5}$	6.78	7.57	11.7
	Age, *yrs*	14.3	14.7	

Direction of variations is not consistent across children nor between FEV_1 and $LCI_{2.5}$, however it is worth noting that substantial changes occur in a very short period of time between evaluations. As these children were not even identified as outliers from the fitted models, they were kept in the longitudinal analyses.

Risk factors for pulmonary exacerbation as recurrent time-to-event data

The majority of children had at least one pulmonary exacerbation (n=207, 84.5%). Mean number of recurrences is 1.8, varying from 0 to 4. First to fourth recurrence of pulmonary exacerbations occurred in 207, 89, 28 and 8 children, respectively (Table 14). Among those with at least one PE, 4% had at most 4 PEs. Therefore, the dataset was truncated after the third event due to small number of events in the fourth stratum.

Table 14. Distribution of anthropometric and clinical markers of CF disease for PE recurrence among 245 children

	Recurrence no.			
	First PE (n=207)	**Second PE (n=89)**	**Third PE (n=28)**	**Fouth PE (n=8)**
Age, *yrs*	11.82 (3.45)	13.14 (2.88)	14.13 (2.17)	14.56 (1.86)
Sex, female	109 (44.5)	41 (46.1)	11 (39.3)	1 (12.5)
BMI, *Z-score*	-0.54 (0.90)	-0.62 (0.87)	-0.83 (0.90)	-0.80 (0.90)
Pancreas Insufficiency	141 (57.6)	59 (66.3)	23 (82.1)	6 (75)
CFTR Genotype				
F508del/other	105 (42.9)	37 (41.6)	11 (39.3)	3 (37.5)
F508del/*F508del*	50 (20.4)	27 (30.3)	10 (35.7)	3 (37.5)
Other/other	90 (36.7)	25 (28.1)	7 (25)	2 (25)
Pseudomonas aeurginosa				
chronic	48 (19.6)	16 (18.0)	7 (25)	2 (25)
free	141 (57.6)	50 (56.2)	18 (64.3)	4 (50)
intermittent	39 (15.9)	22 (24.7)	3 (10.7)	2 (25)
FEV_1, *% predicted*	97.3 (19.6)	96.3 (18.1)	93.8 (16.0)	92.0 (16.5)
FEV_1, Z-*score*	-0.21 (1.61)	-0.29 (1.52)	-0.51 (1.35)	-0.66 (1.39)
$LCI_{2.5}$	10.28 (3.37)	10.63 (3.44)	10.83 (3.44)	11.00 (2.21)
$LCI_{2.5}$, *Z-score*[¶]	3.5 (8.02)	5.42 (8.26)	7.48 (8.8)	11.1 (8.98)

Values are presented as mean (standard deviation) or count (percentage), where not differently expressed. [¶] median (IQR)

Table 14 shows that clinical markers of CF slightly deteriorated as patients developed more recurrences, as one could expect from CF disease.

From the Prentice, Williams and Peterson gap time (PWP-GT) model, $LCI_{2.5}$ and pancreatic status are the two covariates showing evidence of association with the risk of recurrent PE episodes. Children with higher $LCI_{2.5}$ are expected to experience a 6% higher risk of PE recurrence during their follow-up (HR 1.06, 95%CI 1.01 to 1.10), whereas lower risk of recurrent PE episodes is expected in children with pancreatic sufficiency versus those with pancreatic insufficiency (HR 0.59, 95%CI 0.44 to 0.79). Adopting a model with gap-time, this means that children with a previous PE have 26% risk of having a second PE within a year (Figure 23, green line).

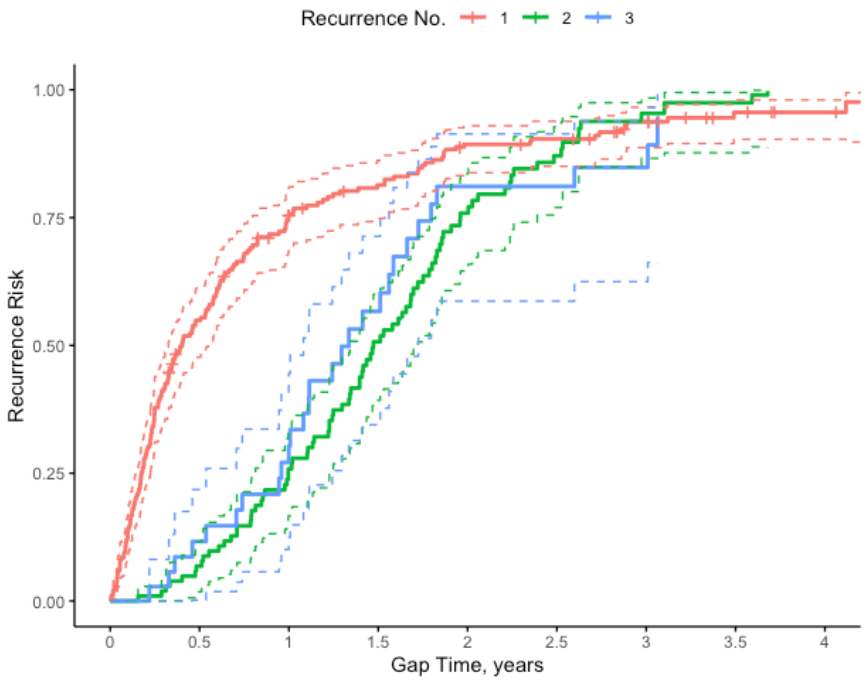

Figure 23. Risk of PE recurrence over time for the first 3 repeated PEs, estimated using gap time among 245 patients with CF. Dotted lines represent 95% confidence intervals estimated using Breslow approximation.

If we consider $LCI_{2.5}$ as dichotomous predictor, as previously defined (see *p.* 22), the above model shows that children with a *normal* LCI have a lower risk of recurrence compared to children with $LCI_{2.5}$ values above ULN (HR 0.72, 95%CI 0.52 to 0.99).

The PWP-GT model adjusted for FEV_1 %predicted (table 15) shows the same association between being pancreatic sufficient and the recurrent PE risk, with a little higher hazard, i.e. 46%. However, lung function measured by FEV_1% predicted is not associated with PE recurrence in this young cohort of children.

Table 15. Results from PWP conditional model using gap time

	Adjusted with $LCI_{2.5}$		Adjusted with FEV_1 % predicted	
	Hazard ratio	P-value	Hazard ratio	P-value
Age	1.00 (0.96;1.04)	0.9740	1.00 (0.97;1.04)	0.8167
Sex				
Male vs Female	0.80 (0.63;1.03)	0.0941	0.86 (0.69;1.08)	0.1893
CFTR Genotype				
F508del/other or Other/other vs F508del/F508del	0.89 (0.68;1.18)	0.4226	0.80 (0.61;1.06)	0.1236
Pancreatic status				
Sufficiency vs insufficiency	0.59 (0.44;0.79)	<0.001	0.54 (0.40;0.72)	<0.001
Pseudomonas aeruginosa infection				
Free vs Intermittent or chronic	0.83 (0.64;1.06)	0.1392	0.78 (0.60;1.00)	0.0500
BMI,Z-score	0.98 (0.86;1.13)	0.8274	0.98 (0.86; 1.13)	0.7937
$LCI_{2.5}$	1.06 (1.01;1.10)	0.0083		
FEV_1, % predicted			1.00 (0.99;1.00)	0.2630

Estimates of association are hazard ratios with 95% confidence intervals. PWP = Prentice, Williams and Peterson

Although not meant for prediction, discriminative ability of the PWP-GT models adjusted for $LCI_{2.5}$ and FEV_1% predicted were respectively 0.58 and 0.59 (adjusted C-Index).

Time-to-pulmonary exacerbations by phenotypes

The same PWP-GT model was also applied to the previous cohort used to assess longitudinal variation over time (*see p.* 49), with the variable *cluster* as the only predictor. Clustering showed evidence of association with PE recurrence (P<0.001), with children in cluster #2 having a lower risk of PE recurrence compared to children defined by cluster #1 (HR 0.46, 95%CI 0.34 to 0.60) (Figure 24).

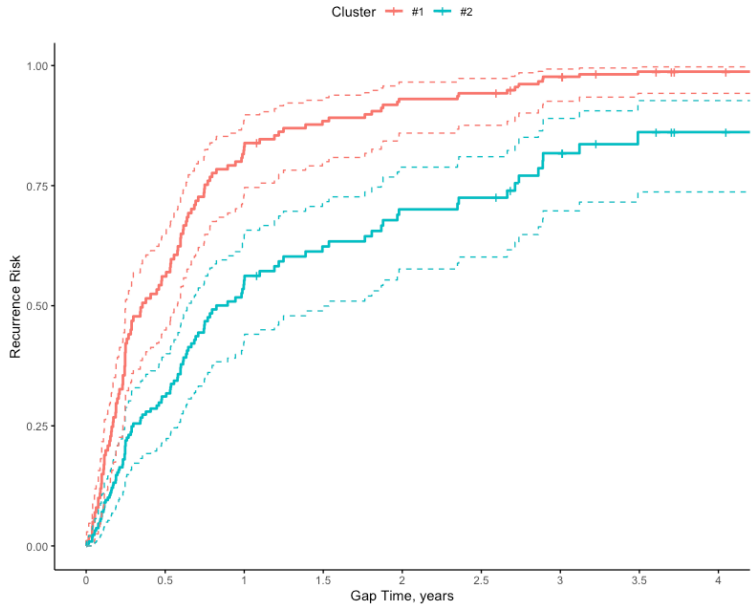

Figure 24. Risk of PE recurrence over time for the first PE, estimated using gap time among 125 patients with CF. Dotted lines represent 95% confidence intervals estimated using Breslow approximation.

Discussion

The aim of this project was investigating the association of some clinical measures used in CF centres to assess CF disease with one marker of lung disease ($LCI_{2.5}$), and to gather information about its behaviour when compared to FEV_1. Cluster analysis was performed to ascertain whether children with CF would express peculiar *trait* in each cluster and within cluster at follow-up, in terms of lung function variation. Eventually, the association between relapses of pulmonary exacerbations and lung function was explored in order to assess the importance of $LCI_{2.5}$ when related to the burden of disease, considering how impactful can be hospitalizations and antibiotic therapies on children and their families.

As theoretically expected, the relationship between FEV_1 and age was better specified when summarized by non-linear fitting. It is well known that pulmonary function varies with age, standing height, sex and ethnicity. Therefore, test results need to be compared to predicted values. New equations from the Global Lung Initiative [64] have been developed, taking into account the changing relationship between FEV_1 and height during the adolescent growth spurt and considering also the discontinuity in the available prediction equation when individuals move from one set of equations to the next. Essentially, GLI provided a linear regression equation with an age-specific correction in the form of the age-spline. The age-correction adopted favoured a better interpretation of the transition from adolescence to adulthood, already describe as non-linear [30]. In the CF Centre of Milano, adolescence is confirmed as the time when children with CF start to express a variable disease expression. FEV_1 %predicted starts decreasing in children around 10 years old, whereas ventilation inhomogeneity keeps increasing, with a linear trend. Using this information merely for descriptive purposes, $MBWN_2$ is able to recognize children younger than approximately 10 years as already exposed to lung disease. Particularly, airflow obstruction of larger airways becomes significantly

impaired after 10 years of age whereas ventilation inhomogeneity characterizes early lung disease in children with CF.

From the cross-sectional results, the present findings show a statistically significant association between age and lung function measures, namely FEV_1 %predicted and $LCI_{2.5}$. Particularly, the FEV_1 increases until 12 years of age whereas $LCI_{2.5}$ steadily increases over time, making FEV_1 poorly sensitive of what occurs in smaller airways until adolescence. Overall, this can be translated into a clinical advantage of using $LCI_{2.5}$ over FEV_1 %predicted alone, suggesting that despite the absence of airflow obstruction, peripheral lung damage causing ventilation inhomogeneity begins earlier. This type of knowledge about the lung clearance index has been extensively discussed in CF literature. [65]

The specific distributions of FEV_1 and $LCI_{2.5}$ challenged the fitting of their relationship with age in the considered sample, hampering the inferential process and the interpretability of the models, as transformations in the forms of square and inverse functions can be difficult to be appreciated. Anyway, the association between $LCI_{2.5}$ and FEV_1 found in this sample reinforce the idea that ventilation inhomogeneity ($LCI_{2.5}$) and air flow obstruction (FEV_1) are two different ways of looking at lung disease in CF. The first refers to the degree of ventilation distribution in the lungs in terms of dis/homogeneity, the latter to the resistance of flow when passing large airways.

Basically, results from regression analyses showed that pancreatic exocrine function, nutritional status and infection by *Pseudomonas aeruginosa* are the strongest predictors of lung function indices in both models. This strengthens the utility of $LCI_{2.5}$ as an early marker of CF disease, especially when children cannot perform spirometry or when FEV_1 show normal/above normal values, keeping in mind that early infection have an impact on future health.

From the present findings, one could say that children are subjected to early small airways modification, to which one must add larger airways remodelling during adolescence. It is worth

mentioning the work by Szczesniak et al., who modelled FEV_1% decline using semi-parametric nonlinear models and suggested that the trends for individuals at risk for prolonged drops in lung function diverged just around 12.9 years of age. [44] One could argue if modifying the linear increasing trend in $LCI_{2.5}$ during young ages could somehow delay the overall impact of small airways on the disease phenotype during adolescence and adulthood.

The importance of adolescence as a key timepoint along the risk of derangement of lung function is also supported by the presence of two already distinct phenotypes at 11 years of age, as reported by the cluster approach used in the present work. Indeed, cluster analysis identified two different profiles of children, interestingly about the same age, equally represented by boys and girls and without statistically or clinically significant difference in nutritional status. Cluster #2 depicts children with less severe genotypes, better lung health and without pancreatic insufficiency. Cluster #1 identifies peers with a more severe expression of disease. The most important thing to consider is that children of same age could show different lung health and a differently compromised lung periphery, only when evaluated by score indices of ventilation inhomogeneity and FEV_1 Z-score together. This is furthermore of interest, taking into consideration that clustering used information derived by $MBWN_2$ and not by spirometry. Although statistically different between the two clusters, FEV_1 % remains above 90% predicted. Again, this reinforces its poor sensitivity in discriminating lung health in children with CF.

The severity of cluster #1 is also supported by the higher prevalence of pulmonary exacerbations compared to cluster #2 and by the absence of hospitalization at all in cluster #2. These children may be identified at elevated risk, being the target of more *personalized* interventions. As commented by Nyilas et al., children with different phenotypes and particularly with distinct ventilation inhomogeneity profiles, may benefit from different therapeutic approach, such as aerosol therapy performed with distinct devices in order to deliver different particle sizes. [47]

Consistency of identified clusters was assessed longitudinally, and two opposite behaviours in terms of FEV_1 %predicted and $LCI_{2.5}$ were detected in the present analyses.

Overall, $LCI_{2.5}$ and FEV_1 did not show evidence of statistical variation over time, both showing a decreasing trend anyway. Considering the time lag between observations, this can be considered a therapeutic success, given that children did not show negative changes in their peripheral lung disease progression, and that these were below the reported variation in literature. [66] Once adjusted for the clinical markers of CF disease and the number of hospitalizations and pulmonary exacerbations experienced in the twelve months before the $MBWN_2$ test, only $LCI_{2.5}$ remains associated with age and with *Pseudomonas aeruginosa*, differently from the study of Davies et al., which did not show any evidence of association between ventilation inhomogeneity and acquisition of *Pseudomonas aeruginosa* in children prior to their pre-school visit. [67] In the present study, the $LCI_{2.5}$ slope was not modified by high-impactful events, such as pulmonary exacerbation or hospitalization, occurred in the year before the first $MBWN_2$ test.

At baseline, children in cluster #2 presented with 3.23 units lower in their $LCI_{2.5}$ compared to children in cluster #1, whereas FEV_1 %predicted was on average 13.3% points higher in cluster #2. Although considering the differences between clusters as statistically and clinically meaningful at each follow-up visit, differences in lung function between clusters did not significantly vary between follow-up visits. Taken all together, results from the present stable cohort of children over a 4-year follow-up show that $LCI_{2.5}$ could be used routinely in the clinic to monitor lung disease, and that it is associated with the detection of *Pseudomonas aeruginosa*, which is known to elicit a pro-inflammatory response in the lung. Most importantly, the present findings show that clusters can characterize children with CF in terms of differences in lung function over time.

In the present work, it remains questionable if the steady trend in $LCI_{2.5}$ should be attributed to any medical decision and subsequent intervention triggered by the evaluation of $LCI_{2.5}$ during the follow-up.

In a recent study, [68] 1-unit larger increase in LCI was not associated with increased antimicrobial use or pathogens load over 2-year follow-up based on LCI-triggered bronchoalveolar lavage. However, the variation in $LCI_{2.5}$ of less than 1 unit seen over 3 follow-up visits in the present work might suggest that smaller variation in $LCI_{2.5}$ may be required to trigger change in the multidisciplinary management of CF, especially in children with stable conditions. Other significative examples come from Stanojevic et al., who showed that $LCI_{2.5}$ significantly deteriorated throughout the 12 month study by a slope of 0.40 (95%CI 0.14 to 0.66) in children aged 3 to 6 years. [40] Also, Perrem et al. showed that $LCI_{2.5}$ increased by 0.87 units from baseline to the symptomatic visit in 98 individuals with CF aged 5 to 17 years. [69] Thus it is plausible that such small variations could be currently used in a clinical setting to guide therapies or require additional assessment. Indeed, it is to be recalled how meaningful was the rescue of MBW test from the armoury of lung function tests in the 2010s in the CF scenario worldwide. It would not be surprising if high $LCI_{2.5}$ values or, at that time, any values above 7 – which was considered a standard cut-off of normality – could have opened to clinicians more options in the care of their CF patients. It is well known that information with a high emotional impact can alter the decisional process, even though the probabilistic rules to guide decisions are already there.

Anyway, the observed stability in ventilation inhomogeneity over a total 4-year follow-up is encouraging for the CF team in Milano, especially considering the recent findings from Sandvik et al., showing that no progression of structural lung disease at CT scan is expected in children with stable $LCI_{2.5}$. [70]

Considering the clinical impact of pulmonary exacerbation on disease progression, as well as the impact of antibiotic therapy on the course of the disease, the last analysis focused on the time to recurrent pulmonary exacerbations from the first $MBWN_2$ test. For this purpose, the event *pulmonary exacerbation* was defined as the moment when children required hospitalization or an antibiotic course, thus representing also mild respiratory events. A recent study [69] showed that $LCI_{2.5}$ worsened with respiratory events such as pulmonary exacerbations in school-age children, and that

recovery was incomplete at follow-up, both in terms of ventilation inhomogeneity and airflow obstruction. This reinforces the idea that lung clearance index is able to track disease progression. Given that recurrence of respiratory events leading to hospitalization or antibiotic therapy may have a significant burden on children and families, we have analysed if $LCI_{2.5}$ could be associated to the risk of PE recurrences. From the present findings, $LCI_{2.5}$ and pancreatic exocrine function are associated the with risk of PE recurrences whereas a separate model with FEV_1 as adjustment covariate did not show any evidence of association between airflow obstruction and the risk of recurrent PEs.

Earlier, Vermeulen et al. [71] showed an association between baseline $LCI_{2.5}$ and FEV_1 and time to first PE in a cohort of 5-19 years old CF patients with CF. Annual PE rate was higher in children with lowest LCI Z-score and FEV_1 Z-score but LCI Z-score was identified as the only predictor of the PE rate in the 12 months following the baseline assessment. Their methods relied on Kaplan-Meier and Negative Binomial regression, which assume that each patient has recurrent events according to individual Poisson event rate which in turn varies according to a specific distribution, i.e. Gamma, across patients. Negative Binomial regression model seems appropriate to estimate recurrent events when information on time is not available, [72] differently from the present study, in which we collected the exact timing of PEs throughout the follow-up. Moreover, the study from Vermeulen et al. used information on PEs up to first event only, potentially leading to an inaccurate evaluation of the association of selected covariates with the event that occurs more than once and that are possibly dependent.

However, both these studies show a relevant association of baseline $LCI_{2.5}$ with the course of CF disease in terms of PE or recurrent PEs, suggesting again how $MBWN_2$ could be useful in the clinic to monitor disease progression. This association is also present in the subset of children assessed longitudinally, showing that a better phenotype is associated with a lower risk of PEs recurrence compared to children phenotypically more compromised.

As regards *CFTR* genotype, it's well known how the association between genotype and the phenotypic expression of the disease is hampered by the influence of environmental and other genetic factors. However, it remains intuitive thinking that patients with mild *CFTR* genotype may show less negative prognostic markers than patients carrying severe genotype. In the present study, a difference distribution of variants was significantly described between clusters, that differentiate one from another also by the number of homozygotic individuals carrying *F508del* (30 *versus* 1). This is the most common mutation worldwide and has long been associated with more severe disease and less favourable clinical outcomes.[73]

In the present cross-sectional analysis, only $LCI_{2.5}$ showed a statistically significant association with genotype, with lower values in patients with different *CFTR* alleles than homozygosis for *F508del*. Very recently, this finding was also reported by Bernasconi et al, showing an association between increased ventilation inhomogeneity and individuals with less *CFTR* function. [74] In our cohort, genotype does not remain longitudinally associated with lung function nor with PE recurrence, suggesting that $LCI_{2.5}$ may be more sensitive than FEV_1 to characterize patients at a specific time, but it is not associated to the type of variants when we look at the impact of genotype on disease progression. For example, McKone et al. reported that the risk of death predicted by the quantitative protein production was higher in patients with severe mutations (class I-III) but not fully explained by lung function measured by FEV_1. [75]

Strengths and limitations

The longitudinal study design reflects usual clinical care, therefore time intervals between visits varied between participants, potentially impacting the observed changes.

$MBWN_2$ is time consuming in little boys and girls aged 5 or younger. Despite research efforts in CF should prioritize children during their *silent* years in order to detect lung function changes, the limited staff at the CF center represents an important barrier to implement this type of lung function

assessment on routine basis in youngsters. For this reason, we were able to include only 2 children under 5 years in the present study.

Future studies should compare these results with regression models taking FEV_1 Z-score as dependent variable, in order to limit possible confounding by age on lung function describing air flow obstruction. Nevertheless, being FEV_1 and $LCI_{2.5}$ two distinct measures of one unique underlying disease, it would be worth exploring their longitudinal variation modelled together, by means of a joint mixed effect models or latent class joint models. In fact, given the existent correlation between $LCI_{2.5}$ and FEV_1, it is unlikely fitting these two covariates at the same time. To account for their intrinsic correlation, such models would give more insights about how these two lung function indices are longitudinally relate to each other.

Conclusion

In light of the present findings, $LCI_{2.5}$ shows interesting associations with clinical characteristics not shared with FEV_1 %predicted, both in cross-sectional and longitudinal analyses. Clustering has shown a disease profile of children who share negative clinical prognostic factors, also in terms of ventilation inhomogeneity. These children are at higher risk of recurrent pulmonary exacerbations as well. Further steps should take into consideration the anatomical correlates of CF lung disease, in order to address to each cluster a specific level of anatomical damage.

Descriptive analysis of the whole CF cohort and clustering approach also support that FEV_1 Z-score has the potential to overcome the limitations reported when describing lung function in terms of % predicted only, especially in younger patients.

Finally, MBW reveals a complementary tool to assess lung function in children with CF, and these results confirm its clinical utility in the evaluation of the course of CF disease. Under such circumstances, the implementation of MBW as part of the routine assessment of individuals with CF by the Regional CF Centre of Milano seems appropriate.

References

1. Sullivan BPO, Freedman SD. Cystic fibrosis. *Lancet* 2009; 373: 1891–1904.
2. Elborn JS. Cystic fibrosis. Lancet 2016. p. 2519–2531.
3. Spoonhower KA, Davis PB. Epidemiology of cystic fibrosis. *Clin. Chest Med.* 2016; 371: 1–8.
4. Zolin A, McKone E, van Rens J. ECFSPR Annual Report. 2014.
5. Rowe SM, Miller S, Sorscher EJ. Cystic Fibrosis. N Engl J Med 2005.
6. Veit G, Avramescu RG, Chiang AN, Houck SA, Cai Z, Peters KW, Hong JS, Pollard HB, Guggino WB, Balch WE, Skach WR, Cutting GR, Frizzell RA, Sheppard DN, Cyr DM, Sorscher EJ, Brodsky JL, Lukacs GL. From CFTR biology toward combinatorial pharmacotherapy: Expanded classification of cystic fibrosis mutations. *Mol. Biol. Cell* 2016; 27: 424–433.
7. Wilschanski M, Zielenski J, Markiewicz D, Tsui LC, Corey M, Levison H, Durie PR. Correlation of sweat chloride concentration with classes of the cystic fibrosis transmembrane conductance regulator gene mutations. *J. Pediatr.* Mosby; 1995; 127: 705–710.
8. Clunes M, Boucher R. The mechanisms of pathogenesis of an inherited lung disorder. *Drug Discov. today Dis. Mech.* 2007; 4: 63–72.
9. Goldman MJ, Anderson GM. Human-Defensin-1 Is a Salt-Sensitive Antibiotic in Lung That Is Inactivated in Cystic Fibrosis. *Cell* 1997; 88: 553–560.
10. Massip-Copiz MM, Santa-Coloma TA. Extracellular pH and lung infections in cystic fibrosis. *Eur. J. Cell Biol.* 2018; 97: 402–410.

11. Colombo C, Costantini D, Rocchi A, Cariani L, Garlaschi ML, Tirelli S, Calori G, Copreni E, Conese M. Cytokine levels in sputum of cystic fibrosis patients before and after antibiotic therapy. *Pediatr. Pulmonol.* 2005; 40: 15–21.

12. Brivio A, Conese M, Gambazza S, Biffi A, Tirelli S, Russo M, Foà M, Colombo C. Pilot Randomized Controlled Trial Evaluating the Effect of Hypertonic Saline With and Without Hyaluronic Acid in Reducing Inflammation in Cystic Fibrosis. *J. Aerosol Med. Pulm. Drug Deliv.* 2016; 29: 1–8.

13. Rottner M, Kunzelmann C, Mergey M, Freyssinet J, Martínez MC. Exaggerated apoptosis and NF-KB activation in pancreatic and tracheal cystic fibrosis cells. *FASEB J.* 2007; 21: 2939–2948.

14. Campodónico VL, Gadjeva M, Paradis-Bleau C, Uluer A, Pier GB. Airway epithelial control of Pseudomonas aeruginosa infection in cystic fibrosis. *Trends Mol. Med.* 2008; 14: 120–133.

15. Moran A, Dunitz J, Nathan B, Saeed A, Holme B, Thomas W. Cystic fibrosis-related diabetes: current trends in prevalence, incidence, and mortality. *Diabetes Care* 2009; 32: 1626–1631.

16. Ooi CY, Durie PR. Cystic fibrosis from the gastroenterologist's perspective. Nat. Rev. Gastroenterol. Hepatol. 2016; p. 175–185.

17. Conway S, Balfour-Lynn IM, De Rijcke K, Drevinek P, Foweraker J, Havermans T, Heijerman H, Lannefors L, Lindblad A, Macek M, Madge S, Moran M, Morrison L, Morton A, Noordhoek J, Sands D, Vertommen A, Peckham D. European cystic fibrosis society standards of care: Framework for the cystic fibrosis centre. *J. Cyst. Fibros.* 2014; 13: S3–S22.

18. Egan ME. Cystic fibrosis transmembrane conductance receptor modulator therapy in cystic fibrosis, an update. *Curr. Opin. Pediatr.* 2020; 32: 384–388.

19. Ramsey BW, Davies J, McElvaney NG, Tullis E, Bell SC, Dřevínek P, Griese M, McKone EF, Wainwright CE, Konstan MW, Moss R, Ratjen F, Sermet-Gaudelus I, Rowe SM, Dong Q, Rodriguez S, Yen K, Ordoñez C, Elborn JS. A CFTR Potentiator in Patients with Cystic Fibrosis and the G551D Mutation. *N. Engl. J. Med.* 2011; 365: 1663–1672.

20. Nick JA, St. Clair C, Jones MC, Lan L, Higgins M. Ivacaftor in cystic fibrosis with residual function: Lung function results from an N-of-1 study. *J. Cyst. Fibros.* 2020; 19: 91–98.

21. Wainwright CE, Elborn JS, Ramsey BW, Marigowda G, Huang X, Cipolli M, Colombo C, Davies JC, De Boeck K, Flume PA, Konstan MW, McColley SA, McCoy K, McKone EF, Munck A, Ratjen F, Rowe SM, Waltz D, Boyle MP. Lumacaftor–Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. *N. Engl. J. Med.* 2015; 373: 220–231.

22. Taylor-Cousar JL, Munck A, McKone EF, van der Ent CK, Moeller A, Simard C, Wang LT, Ingenito EP, McKee C, Lu Y, Lekstrom-Himes J, Elborn JS. Tezacaftor–Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del. *N. Engl. J. Med.* 2017; 377: 2013–2023.

23. Heijerman HGM, McKone EF, Downey DG, Van Braeckel E, Rowe SM, Tullis E, Mall MA, Welter JJ, Ramsey BW, McKee CM, Marigowda G, Moskowitz SM, Waltz D, Sosnay PR, Simard C, Ahluwalia N, Xuan F, Zhang Y, Taylor-Cousar JL, McCoy KS, McCoy K, Donaldson S, Walker S, Chmiel J, Rubenstein R, Froh DK, Neuringer I, Jain M, Moffett K, Taylor-Cousar JL, 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.

24. Harness-Brumley C, Elliott A, Rosenbluth D, Raghavan D, Jain R. Gender differences in outcomes of patients with cystic fibrosis. *J. women's Heal.* 2014; : 1012–1020.

25. De Boer K, Vandemheen KL, Tullis E, Doucette S, Fergusson D, Freitag A, Paterson N, Jackson M, Lougheed MD, Kumar V, Aaron SD. Exacerbation frequency and clinical outcomes in adult patients with cystic fibrosis. *Thorax* 2011; 66: 680–685.

26. Sawicki GS, Rasouliyan L, McMullen AH, Wagener JS, McColley SA, Pasta DJ, Quittner AL. Longitudinal assessment of health-related quality of life in an observational cohort of patients with cystic fibrosis. *Pediatr. Pulmonol.* 2011; 46: 36–44.

27. Sanders DB, Bittner RCL, Rosenfeld M, Hoffman LR, Redding GJ, Goss CH. Failure to recover to baseline pulmonary function after cystic fibrosis pulmonary exacerbation. *Am. J. Respir. Crit. Care Med.* 2010; 182: 627–632.

28. Briesacher BA, Quittner AL, Fouayzi H, Zhang J, Swensen A. Nationwide trends in the medical care costs of privately insured patients with cystic fibrosis (CF), 2001-2007. *Pediatr. Pulmonol.* 2011; 46: 770–776.

29. Flume P a., Mogayzel PJ, Robinson K a., Rosenblatt RL, Quittell L, Marshall BC, Cunningham J, Downs A, Fleige J, Goss C, Gutierrez H, Hazle L, Kuhn R, Lester M, Sabadosa K, Vender RL, White TB, Willey-Courand DB, Chin M, Milla CE, Wagener J, Oermann CM, Ferkol T, Hoover W, Howenstine MS, Rock MJ, Retsch-Bogart G, Rubenstein R, Schechter M, Orenstein DM, et al. Cystic fibrosis pulmonary guidelines: Pulmonary complications: Hemoptysis and pneumothorax. *Am. J. Respir. Crit. Care Med.* 2010; 182: 298–306.

30. Harun SN, Wainwright C, Klein K, Hennig S. A systematic review of studies examining the rate of lung function decline in patients with cystic fibrosis. *Paediatr. Respir. Rev.* 2016; 20: 55–56.

31. Keogh RH, Szczesniak R, Taylor-Robinson D, Bilton D. Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: A longitudinal study using UK patient registry data. *J. Cyst. Fibros.* 2018; 17: 218–227.

32. Tiddens H a WM, Donaldson SH, Rosenfeld M, Paré PD. Cystic fibrosis lung disease starts in the small airways: Can we treat it more effectively? *Pediatr. Pulmonol.* 2010; 45: 107–117.

33. Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, Thamrin C, Arets HGM, Aurora P, Fuchs SI, King GG, Lum S, Macleod K, Paiva M, Pillow JJ, Ranganathan S, Ratjen F, Singer F, Sonnappa S, Stocks J, Subbarao P, Thompson BR, Gustafsson PM. Consensus statement for inert gas washout measurement using multiple- and single-breath tests. *Eur Respir J* 2013; 41: 507–522.

34. Horsley A. Lung clearance index in the assessment of airways disease. *Respir. Med.* 2009; 103: 793–799.

35. Saunders C, Bayfield K, Irving S, Short C, Bush A, Davies JC. Developments in multiple breath washout testing in children with cystic fibrosis. *Curr. Med. Res. Opin.* 2017; 33: 613–620.

36. Usemann J, Yammine S, Singer F, Latzin P. Inert gas washout: background and application in various lung diseases. *Swiss Med. Wkly.* 2017; 147: 1–13.

37. Anagnostopoulou P, Latzin P, Jensen R, Stahl M, Harper A, Yammine S, Usemann J, Foong RE, Spycher B, Hall GL, Singer F, Stanojevic S, Mall MA, Ratjen F, Ramsey KA. Normative data for multiple breath washout outcomes in school-aged Caucasian children. *Eur. Respir. J.* 2020; 55

38. Houltz B, Green K, Lindblad A, Singer F, Robinson P, Nielsen K, Gustafsson P. Tidal N2 washout ventilation inhomogeneity indices in a reference population aged 7-70 years. *Eur. Respir. Soc. Conf.* 2012.

39. Lum S, Stocks J, Stanojevic S, Wade A, Robinson P, Gustafsson P, Brown M, Aurora P, Subbarao P, Hooe AF, Sonnappa S. Age and height dependence of lung clearance index and functional residual capacity. *Eur. Respir. J.* 2013; 41: 1371–1377.

40. Stanojevic S, Davis SD, Retsch-Bogart G, Webster H, Davis M, Johnson RC, Jensen R, Pizarro ME, Kane M, Clem CC, Schornick L, Subbarao P, Ratjen FA. Progression of lung disease in preschool patients with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* 2017; 195: 1216–1225.

41. Kraemer R, Blum A, Schibler A, Ammann RA, Gallati S. Ventilation Inhomogeneities in Relation to Standard Lung Function in Patients with Cystic Fibrosis. *Am J Respir Crit Care Med* 2005; 171: 371–378.

42. Horsley ALEX, Siddiqui SA. Putting lung function and physiology into perspective : cystic fibrosis in adults. *Respirology* 2015; : 33–45.

43. Oude Engberink E, Ratjen F, Davis SD, Retsch-Bogart G, Amin R, Stanojevic S. Inter-test reproducibility of the lung clearance index measured by multiple breath washout. *Eur. Respir. J.* 2017; 50: 1700433.

44. Szczesniak RD, Li D, Su W, Brokamp C, Pestian J, Seid M, Clancy JP. Phenotypes of rapid cystic fibrosis lung disease progression during adolescence and young adulthood. *Am. J. Respir. Crit. Care Med.* 2017; 196: 471–478.

45. Conrad DJ, Bailey BA. Multidimensional clinical phenotyping of an adult cystic fibrosis patient population. *PLoS One* 2015; 10: 1–14.

46. Hafen GM, Hurst C, Yearwood J, Smith J, Dzalilov Z, Robinson PJ. A new scoring system in Cystic Fibrosis: Statistical tools for database analysis - A preliminary report. *BMC Med. Inform. Decis. Mak.* 2008; 8: 1–11.

47. Nyilas S, Singer F, Kumar N, Yammine S, Meier-Girard D, Koerner-Rettberg C, Casaulta C, Frey U, Latzin P. Physiological phenotyping of pediatric chronic obstructive airway diseases. *J. Appl. Physiol.* 2016; 121: 324–332.

48. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates a., Crapo R, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntrye N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wagner J. Standardisation of spirometry. *Eur. Respir. J.* 2005; 26: 319–338.

49. Graham BL, Steenbruggen I, Barjaktarevic IZ, Cooper BG, Hall GL, Hallstrand TS, Kaminsky DA, McCarthy K, McCormack MC, Miller MR, Oropez CE, Rosenfeld M, Stanojevic S, Swanney MP, Thompson BR. Standardization of spirometry 2019 update an official American Thoracic Society and European Respiratory Society technical statement. *Am. J. Respir. Crit. Care Med.* 2019; 200: E70–E88.

50. Jensen R, Green K, Gustafsson P, Latzin P, Pittman J, Ratjen F, Robinson P, Singer F, Stanojevic S, Yammine S. Standard Operating Procedure: Multiple Breath Nitrogen Washout 2013. Available from: http://www.mbwtraining.com/ECFS_MBW_SOP.pdf.

51. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MSMM, Zheng J, Stocks J, Schindler C. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.

52. Subbarao P, Milla C, Aurora P, Davies JC, Davis SD, Hall GL, Heltshe S, Latzin P, Lindblad A, Pittman JE, Robinson PD, Rosenfeld M, Singer F, Starner TD, Ratjen F, Morgan W. Multiple-Breath Washout as a Lung Function Test in Cystic Fibrosis A Cystic Fibrosis Foundation Workshop Report. *Am J Respir Crit Care Med* 2015; 12.

53. Lombardi E, Gambazza S, Pradal U, Braggion C. Lung clearance index in subjects with cystic fibrosis in Italy. *Ital. J. Pediatr.* 2019;45(1):56

54. Robinson PD, Latzin P, Ramsey KA, Stanojevic S, Aurora P, Davis SD, Gappa M, Hall GL, Horsley A, Jensen R, Lum S, Milla C, Nielsen KG, Pittman JE, Rosenfeld M, Singer F, Subbarao P, Gustafsson PM, Ratjen F. Preschool Multiple-Breath Washout Testing. An Official American Thoracic Society Technical Statement. *Am. J. Respir. Crit. Care Med.* 2018; 197: e1–e19.

55. van Buuren S. Flexible Imputation of Missing Data. Second Edi. CRC Press.

56. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Med. Res. Methodol.* 2009; 9: 1–8.

57. Gower J. A general coefficient of similarity and some of its properties. *Biometrics* 1971; 27: 857–871.

58. Kassambara A. Practical guide to Cluster Analysis in R. STHDA; 2017.

59. Rousseeuw P. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 1987; 20: 53–65.

60. Prentice RL, Williams BJ, Peterson A V. On the Regression Analysis of Multivariate Failure Time Data. *Biometrika* 1981; 68: 373–379.

61. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2019.

62. Zolin A, Orenti A, Naehrlich L, van Rens J. ECFSPR Annual Report 2017. 2019.

63. Stanojevic S, Bilton D. Global Lung Function Initiative equations improve interpretation of FEV 1 decline among patients with cystic fibrosis. *ERJ* 2015; 10–12.

64. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MSM, Zheng J. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.

65. Short C, Saunders C, Davies JC. Utility of lung clearance index in CF: What we know, what we don't know and musings on how to bridge the gap. *J. Cyst. Fibros.* 2020; 3–6.

66. Liou TG, Elkin EP, Pasta DJ, Jacobs JR, Konstan MW, Morgan WJ, Wagener JS. Year-to-year changes in lung function in individuals with cystic fibrosis. *J. Cyst. Fibros.* 2010; 9: 250–256.

67. Davies G, Stanojevic S, Raywood E, Duncan JA, Stocks J, Lum S, Bush A, Viviani L, Wade A, Calder A, Owens CM, Goubau C, Carr SB, Bossley CJ, Pao C, Aurora P. An observational study of the lung clearance index throughout childhood in cystic fibrosis: early years matter. *Eur. Respir. J.* 2020; 56.

68. Voldby C, Green K, Kongstad T, Ring AM, Sandvik RM, Skov M, Buchvald F, Pressler T, Nielsen KG. Lung clearance index-triggered intervention in children with cystic fibrosis – A randomised pilot study. *J. Cyst. Fibros.* Elsevier B.V.; 2020; .

69. Perrem L, Stanojevic S, Shaw M, Jensen R, McDonald N, Isaac SM, Davis M, Clem C, Guido J, Jara S, France L, Solomon M, Grasemann H, Waters V, Sweezey N, Sanders DB, Davis SD, Ratjen F. Lung Clearance Index to Track Acute Respiratory Events in School-age Children with Cystic Fibrosis. *Am. J. Respir. Crit. Care Med.* 2020; : 1–5.

70. Sandvik RM, Kongstad T, Green K, Voldby C, Buchvald F, Skov M, Pressler T, Nielsen KG. Prospective longitudinal association between repeated multiple breath washout measurements and computed tomography scores in children with cystic fibrosis. *J. Cyst. Fibros.* 2020;4:2121.

71. Vermeulen F, Proesmans M, Boon M, Havermans T, Boeck K De. Lung clearance index predicts pulmonary exacerbations in young patients with cystic fibrosis. *Thorax* 2013; 1–4.

72. Yadav CP, Sreenivas V, Khan M, Pandey R. An Overview of Statistical Models for Recurrent Events Analysis: A Review. *Epidemiol.* 2018; 8.

73. Salvatore D, Buzzetti R, Baldo E, Forneris MP, Lucidi V, Manunza D, Marinelli I, Messore B, Neri AS, Raia V, Furnari ML, Mastella G. An overview of international literature from cystic fibrosis registries. Part 3. Disease incidence, genotype/phenotype correlation, microbiology, pregnancy, clinical complications, lung transplantation, and miscellanea. *J. Cyst. Fibros.* 2011; 10: 71–85.

74. Bernasconi N, Kieninger E, Shaw M, Kurz J, Moeller A, Ratjen F, Rochat I, Stanojevic S, Singer F. CFTR-function and ventilation inhomogeneity in individuals with cystic fibrosis. *J. Cyst. Fibros.* 2020; 18:S1569-1993(20)30941-3.

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