Chest computed tomography; a validated surrogate endpoint of cystic fibrosis lung disease?

CT; a validated endpoint in cystic fibrosis?

Martine Loeve¹,², Gabriel P. Krestin², Margaret Rosenfeld³, Marleen de Bruijne⁴,⁵, Stephen M. Stick⁶, Harm A. Tiddens¹,²

¹Department of Pediatric Pulmonology and Allergology, Erasmus Medical Center-Sophia Children’s Hospital, The Netherlands
²Department of Radiology, Erasmus Medical Center, The Netherlands
³Division of Pulmonary Medicine, Department of Pediatrics, University of Washington School of Medicine, Seattle, Washington, USA
⁴Biomedical Imaging group Rotterdam, Departments of Radiology & Medical informatics, Erasmus MC, Rotterdam, the Netherlands
⁵Department of Computer Science, University of Copenhagen, Denmark
⁶Division of Pulmonary Medicine, Princess Margaret Hospital for Children, Perth, Australia

Prof. Dr. H.A.W.M. Tiddens, MD, PhD
Dr. Molewaterplein 60
3015 GJ Rotterdam
The Netherlands
0031 10 7036263
0031 10 7036800
h.tiddens@erasusmc.nl

This study was supported by grants from the Sophia CF research fund, the Dutch Cystic Fibrosis Foundation (NCFS), and the Italian CF Fund (IERFC). None of the sponsors were involved in the study design, data collection, analysis, interpretation of the data, writing of the report, or in the decision to submit the paper for publication.

Review
Word count
5734 words

Number of figures and tables
1 Figure, 1 Table
ABSTRACT
Clinical trials for the treatment of cystic fibrosis (CF) lung disease are important to test and optimize new therapeutic interventions. To evaluate the effect of these interventions, sensitive and accurate outcome measures are needed. The most commonly used endpoints are spirometric variables such as the forced expiratory volume in one second (FEV₁) and respiratory tract exacerbations (RTE). Unfortunately, these endpoints are relatively insensitive to monitor progression of CF lung disease, and thus require a large number of patients when used in clinical studies. In addition, these endpoints are not suitable to study CF lung disease in young children. Chest computed tomography (CT) holds great promise for use as a sensitive surrogate endpoint in CF. A large body of evidence has been produced to validate the use of chest CT as primary endpoint to study CF lung disease. However, before chest CT can be used in clinical trials, it has to be recognized as a validated surrogate endpoint by regulatory agencies. The aim of this review is to summarize what is currently known about the use of chest CT as surrogate endpoint in clinical trials in CF.

KEYWORDS
Clinical trials
Monitoring lung disease
Introduction

Cystic fibrosis (CF) lung disease is the primary cause of death in most CF patients. Thanks to improved treatment, median survival currently approaches 38 years.(1) New therapeutic agents such as inhaled antibiotics and disease modifying agents are in development to further improve treatment.(2-4) To test the effectiveness and safety of these drugs in randomized clinical trials, sensitive and accurate outcome measures are needed.(5) Before an outcome measure can be used as a primary endpoint in a clinical study, it has to meet stringent requirements. First, it must reflect the presence and severity of the disease. Second, it has to reflect clinically meaningful improvement in disease severity when the disease is treated effectively. Third, the measurement must be reproducible. Fourth, changes in the endpoint should closely match changes in the true outcome.(6) Two types of endpoints can be distinguished; clinical or “true” endpoints, and surrogate endpoints. The US Food and Drug Administration (FDA) defines clinical endpoints as direct measures of how a patient feels, functions, or survives, and are thus expected to predict the effect of a therapy. The FDA defines surrogate endpoints as laboratory measurements or physical signs that can be used as a substitute for a clinical endpoint.(7) Surrogate endpoints are generally used when the clinical benefit may not be detectable in trials of reasonable cost, duration, or size.(8) FDA regulations state that a surrogate endpoint is considered to be “reasonably likely to predict clinical benefit and, therefore, useable for drug approval if there is evidence based on
epidemiologic, therapeutic, pathophysiologic, or other data supporting the association of the surrogate with the clinical benefit". (8)

In CF, two clinical endpoints are currently used; respiratory tract exacerbation rate (RTE-R) and quality of life (QoL). (9) RTE-R is an important endpoint that has been shown to increase with age and more severe lung function impairment. (10) In addition, there is a clear association between RTE-R and survival in CF. (11) Unfortunately, the use of RTE-R has disadvantages. It is a relatively insensitive endpoint, especially in (young) children with early lung disease, who generally experience less frequent exacerbations compared with older children. As a result, RTE-R requires a large sample size when used in clinical studies. (12) Furthermore, there is no universal consensus regarding the definition of an RTE. (13)

QoL is another important clinical endpoint. The FDA accepts patient reported outcome measures as primary or secondary endpoints if they are appropriate for the disease, product, and indication. (14) QoL has the advantage of measuring directly how a patient reports to feel or function. The two most commonly used CF-specific QoL questionnaires (the CF questionnaire-revised (CFQ-R), and the CF Quality of Life Questionnaire) are well validated with demonstrated reliability, validity, and sensitivity. (15-16) The CFQ-R has been used as a primary endpoint in a study evaluating the effectiveness of an inhaled antibiotic against Pseudomonas Aeruginosa. (17) However, to date there is still a lack of systematically validated QoL questionnaires in young children. In addition, further work is needed to standardize the use of QoL instruments in clinical research. Furthermore, the development of new QoL instruments that are sensitive to
smaller and earlier changes in symptoms is needed, as improvements in overall health and lung function are seen in successive CF birth cohorts. (14)

The most commonly used surrogate endpoint in CF is the forced expiratory volume in one second (FEV₁). The FEV₁ is indirectly related to structural lung damage. When structural damage affects large lung volumes, low values are observed. (18) Thanks to improvements in CF therapy over the last decades, FEV₁ has become relatively insensitive for monitoring CF lung disease progression. Currently, FEV₁ is in the normal range for most patients until adolescence, and average annual FEV₁ decline is less than 1%. (19) In addition, FEV₁ is insensitive to detect early and localized structural changes (18), is difficult for young children to perform, and is not suitable for infants and most pre-school children. (9) Thus, there is a need for new, more sensitive surrogate endpoints in CF that reflect mild lung disease and that can be used to assess lung disease in infants and young children.

A promising endpoint that has been extensively studied since the mid nineties is chest computed tomography (CT) images of the lungs. Importantly, using chest CT, structural abnormalities can be easily observed. Various methods have been developed to quantify these structural abnormalities. In addition, a large number of studies have been done to validate chest CT as a surrogate endpoint. The aim of this review is to summarize what is currently known about the use of chest CT as a surrogate endpoint in clinical trials in CF. In addition, we will discuss its advantages and disadvantages, and the future work needed to further improve sensitivity and accuracy of chest CT for use in clinical trials.
CF lung disease

At birth, CF patients have macroscopically normal lungs. The structural changes related to CF however, occur early. Pathology studies have shown the presence of structural abnormalities such as bronchiectasis, and mucus impaction even in infants with CF aged 0 to 4 months. In addition, all children in this age group were found to have evidence of bronchial wall inflammation. Airway wall thickening, a frequently observed abnormality in CF, has been correlated with airway inflammation. In other lung diseases, such as chronic obstructive pulmonary disease (COPD) and asthma, the severity of airflow obstruction is proportional to the severity of airway wall thickening. Similar correlations have been reported in CF. Furthermore, a 3-fold increase in airway wall thickening was found in specimens from CF patients compared with specimens from COPD patients. In addition, destruction of the epithelial layer and a substantial loss of cartilage have been described in CF. All these factors are likely to contribute to the airflow obstruction that is present in CF. To monitor CF-related lung changes in vivo, CT scanning can be used, as it was shown to correlate well to pathologic findings such as bronchiectasis.

Image analysis

An important condition for the use and validation of chest CT as surrogate endpoint is that the structural abnormalities can be quantified in a reproducible
manner. This can be done using manual (semi-)quantitative scoring systems and/or automated systems.

*Manual scoring systems:* For CT scoring, a range of systems is available with proven reproducibility. In a comparative study, within and between reader reproducibility for 5 scoring systems was found to be good with intraclass correlation coefficients of 0.74 and higher. These semi-quantitative scoring systems evaluate structural abnormalities by assigning scores ranging from 0 to 3 severity and/or extent. These systems were mainly developed to quantify abnormalities in mild to moderate CF lung disease. For CF patients with severe advanced lung disease (SALD), a dedicated SALD scoring system was developed. This system uses a digital grid to annotate all lung tissue into 4 components: (1) infection/inflammation (including several abnormalities, from which bronchiectasis is the most important one), (2) trapped air/hypoperfusion, (3) bulla/cysts, and 4) normal/hyperperfusion. The SALD system has been shown to correlate with the Brody-II system developed for mild disease, with good reproducibility between and within observers.

*Automated systems:* Ideally, automated systems should be used for CT quantification, as they allow more rapid analyses with near perfect intratest reproducibility. Various semi-automated (30-35) and fully automated (36-43) systems have been developed. The semi-automated systems have been shown capable of measuring airway wall thickness, bronchial diameter (30-31), and trapped air. These systems have been validated in CF, and could differentiate patients from controls. Furthermore, parameters have been
correlated with visual scores from radiologists (30-31), and pulmonary function
tests (PFTs). (30-33) For longitudinal follow up, progression of trapped air has
been tracked. In a study by Loeve et al., two routine expiratory CTs over 2 years
were matched, and the proportion of stable, disappeared and new trapped air
was measured. With this novel approach, trapped air was found to have a stable
component. (34) Quantitative trapped air measurements have been used in
intervention trials. (35) In a 1-year randomized placebo-controlled trial with
dornase alpha, 25 children with mild CF were enrolled. The results showed that
quantitative trapped air measures could discriminate differences in treatment
effects in children with mild CF at 1 of the 2 study visits. The disadvantages of
semi-automated systems is that they require user input to mark the trachea (31),
airway wall and vessel (30) or thresholds for trapped air. (32-34) This in contrast
to fully automated systems that require no user input. Currently available
automated systems are able to quantify the bronchial tree (40-42) and measure
airway wall thickness and bronchial diameter. (36, 38-39, 43-44) In addition, a
system was proposed for matching and labelling of airway trees to allow within
and between patient comparisons. (41) These automated systems were validated
using CT images from lung cancer screening trials (38), smokers (39), healthy
subjects (39, 42), a random set of cases (37, 44), asthma patients (43), or
phantoms. (39-40, 42) Measurements correlated well with PFTs (38) and manual
expert readings. (36-37, 41) Unfortunately, validation in CF is lacking in all but
one study. (40) In addition, there are still substantial technical issues to be
resolved. For example, the influence on the measurements by factors such as
scan parameters, inflation level, effect of visual input on reproducibility for the semi-automated systems and clinical validity of the measurements.\footnote{45} Thus, automated systems are promising, but further development is needed before these systems can be used in clinical trials.

**Presence and severity of disease**

The first validation requirement for endpoints is that it must reflect the presence and severity of the disease. Currently, chest CT is considered the gold standard for detection of bronchiectasis.\footnote{46} An important advantage of CT is that it easily allows recognition of the structural lung abnormalities characteristic for CF, which are not present in normal subjects. Furthermore, CT can detect abnormalities early. In cohort studies in children with CF diagnosed through newborn screening structural abnormalities such as bronchiectasis, trapped air, and mucus impaction could be observed even in asymptomatic infants.\footnote{47-50} In addition, bronchiectasis in infants and older children was shown to be progressive.\footnote{47, 51-52, 53} In children with CF diagnosed through newborn screening, bronchiectasis was present on chest CT in around 10\% in the first year of life. By the age of five years over 50\% of children had bronchiectasis.\footnote{Stick, 2009 \#268, 54} Trapped air on expiratory CT was observed in two thirds of children with CF in the first year of life and continued to persist in around two thirds in the first five years of life.\footnote{47, 53-54} Furthermore, CT proved to be more sensitive in detecting and monitoring bronchiectasis and other structural abnormalities than PFT-related parameters.
Finally, bronchiectasis and trapped air are the most important components of end stage CF lung disease. Other structural abnormalities that can be observed on chest CT in CF are airway wall thickening, consolidations and bulla/cysts. An important advantage of CT over PFTs is that structural lung abnormalities can be differentiated in the above described components.

To summarize, there is a large body of data supporting the concept that CT is sensitive to detect the presence and monitor the severity of structural changes relevant for patients with CF starting in infancy into adulthood.

Response to treatment

The second validation requirement for endpoints is that it has to show improvement upon successful treatment of the disease. CT scores have been shown to improve with antibiotic treatment for an RTE. Shah et al. studied 27 adult CF patients, and showed that CT scores for airway wall thickening, mucus, and air fluid levels in bronchiectasis and centrilobular nodules improved in respectively 2/19 (11%), 6/18 (33%), 2/2 (100%) and 4/11 (36%) of patients in response to treatment. Similar findings were reported for pediatric patients. Robinson et al. studied 17 pediatric and adult CF patients (mean age 17.2 years), and showed that CT total and mucus score was significantly reduced after treatment for an RTE. Brody et al. studied 8 pediatric CF patients (mean age 12.7 years) experiencing a total of 15 RTEs, and showed that CT airway wall
thickening and mucus scores significantly improved after treatment. (57) Even in very young children, these improvements were observed. Davis et al. studied 13 young children with CF (mean age 17 months), and showed that CT total, combined bronchiectasis-bronchial dilatation and hyperinflation scores significantly reduced after treatment. (58)

CT score improvement after treatment with dornase alpha and tobramycin solution for inhalation is less well established. Robinson et al. studied 25 children with CF, and showed that quantitative trapped air score and mean total mucus score decreased significantly in the dornase alfa group compared with placebo. However, these effects were only significant at 1 of the 2 study visits. (35) Using a different scoring system, these CT scores could discriminate between treatment and placebo group at 1 of the 2 study visits, but only when combined with PFTs as a composite score. (59) Modified scores were also used in other studies. Studying dornase alpha, Nasr et al. reported significant increases in CT scores in the treatment group versus placebo. However, this was only present when CT scores were calculated as a "gain score". (60) Studying tobramycin solution for inhalation, they reported a trend towards improvement after 1 cycle of treatment using a modified total CT score. (61) After 3 treatment cycles, they found a significant decrease in airway wall thickening score in the treatment group versus placebo. (62) Bronchiectasis is an important outcome parameter. Since bronchiectasis is defined as an irreversible structural change, intervention studies can investigate whether the development of bronchiectasis can be prevented or whether its progression can be slowed down. All intervention studies cited above
included bronchiectasis as an outcome measure. Surprisingly, despite the irreversible nature of bronchiectasis, decreased bronchiectasis scores were observed in the intervention group in all but 1 study.(35, 59-62) However, whether the observed reductions in bronchiectasis score were statistically significant was not stated. Most likely, these reductions can be explained by the CT protocols used. Most of the above mentioned studies used limited slice protocols. It is well recognized that the variability of these protocols to detect bronchiectasis is higher and its sensitivity lower relative to volumetric CT protocols.(63)

Thus, while other CT scores can improve with treatment, the effect of treatment on bronchiectasis needs to be further studied. Until the efficacy of antibiotics, mucolytics or disease modifiers on the development of bronchiectasis has been reported, this requirement for bronchiectasis score has not yet been completely met.

**Standardization**

The third validation requirement for endpoints is reproducibility of the measurement. For CT, it is important to standardize the scanning protocol to optimize reproducibility. The most important variable that requires standardization is the lung volume during scanning. Traditionally, in patients aged 5 years and above voluntary breath holding is used during scanning. Patients are instructed by the CT technician or a recorded voice to perform a maximal inhalation manoeuvre, and then hold their breath during scanning to obtain end-inspiratory
images. Next, patients are asked to maximally exhale and hold their breath to obtain end-expiratory images. With this technique however, one cannot assess whether a patient was able to follow the instruction correctly and hold his breath at the maximum in- and exhalation level. For many patients and specifically for children, these manoeuvres are difficult to perform. In a pediatric study (mean age 12 years), voluntary breath hold scan volumes were compared with plethysmographic lung volume measurements prior to scanning. Average inspiratory volume was acceptable at 77% of total lung capacity (TLC), but ranged from 55% to 106%. Average expiratory volume was 86% of functional residual capacity (FRC) and 140% of residual volume (RV), suggesting that expiratory manoeuvres are even more difficult for children.\(^{64}\) Thus, there is a need for better standardization of breath holding during CT scanning. This can be obtained by using a spirometer combined with instructions by a lung function technician. Prior to scanning, the patient practices the breathing manoeuvres with the spirometer coached by the technician. Next, the patient performs these manoeuvres during scanning, again coached by the lung function technician.\(^{65}\) Scanning will commence when the required volumes are obtained by the patient. The aim of this procedure is to obtain an inflation level near 95% of slow vital capacity (SVC), and an expiration level near 5-12% of SVC \(^{48,66}\) Recently, the volume-monitored technique has been described. This method combines coaching with a portable spirometer that generates respiratory tracings to aid the patient in achieving targeted lung volumes.\(^{67}\) With this method, patients as
young as 4 years were able to achieve reproducible images at 95% of full inflation and at 77% of vital capacity for the expiratory images.

In very young children, lung volume standardization requires a different approach. For these children, a non-invasive pressure-controlled ventilation (PCV) technique under general anaesthesia or sedation has been introduced. This technique requires no intubation and starts off by hyperventilating the child by giving a series of augmented breaths using positive pressure applied via a facemask to induce a physiologic respiratory pause. During this pause, the lungs are imaged at full inflation by maintaining a positive facemask pressure (25–30 cm water) for inspiratory images, and resting end-exhalation by applying no mask pressure for expiratory images. Both spirometer and PCV techniques have been shown to be highly reproducible.

Volume control is particularly important for assessment of bronchiectasis and trapped air. In a pediatric study, bronchiectasis was identified on 30% of images obtained at end-inspiration using PCV techniques compared with 6% of images obtained during quiet breathing. Trapped air was seen in 45% of images obtained at end-expiration compared with 19% of images obtained during quiet breathing. Thus, for standardization of chest CT, volume control is an important condition for both inspiratory and expiratory scans.
Correlation with true endpoints

The fourth validation requirement for endpoints is that changes in the endpoint should closely match changes in the true outcome. The first argument providing that CT meets this criterion is that higher CT scores have been correlated with characteristics associated with more severe disease. For example, higher CT scores were reported for patients with pancreatic insufficiency compared with patients with sufficient pancreatic function.(70) In addition, CT scores strongly correlated with Pseudomonas Aeruginosa acquisition, a well established risk factor for progressive CF lung disease.(71-73) The second argument that CT meets this criterion, is that CT scores have been linked to the true endpoints RTE-R, QoL, and survival.

RTE-R: Two studies have shown a correlation between severity of structural changes on CT and RTE-R.(74-75) The first study described 61 subjects from the Pulmozyme Early Intervention Trial, including 6- to 10-year-old children with well preserved lung function (forced vital capacity ≥ 85%).(74) CTs and PFTs were performed at the beginning and end of the 2-yr trial during which RTE-R was recorded. Nine out of 61 subjects experienced a total of 22 RTEs. PFTs and CT scores at baseline correlated significantly with RTE-R ($r_s$ from -0.40 to 0.30), although none of these variables predicted RTEs with high accuracy.(74) The second study was performed in an unselected cohort of 115 children and adolescents with CF.(75) In this retrospective study, routine chest CTs and PFTs performed during annual checkup were collected with two years’ follow up in which RTE-R was recorded. Fifty-one subjects experienced a total of 148 RTEs.
Bronchiectasis was found to be the strongest predictor of RTEs of all CT parameters, and added significantly to the predictive value of FEV₁.

QoL: A recent study has shown an association between bronchiectasis and impaired QoL. (76) In this cross-sectional study, the effect of CT scores on QoL was assessed using routine CTs and CFQ-Rs from 72 children and adolescents with CF. CTs were scored using a modified Brody-II scoring system. Significant correlations between the respiratory domain and bronchiectasis and trapped air scores were found ($r_s$ -0.39 and -0.30, respectively).

Survival: Recently, a correlation between CT and survival has been established. In a multi-center study including 366 CF patients awaiting lung transplantation, routine CT scans acquired at the time of screening were scored with the (semi-automated) SALD scoring system. (29) SALD infection/inflammation score (including bronchiectasis) was significantly correlated with waiting list survival, and added to the predictive value of the lung allocation score. (77) The correlation with survival is an important addition to the portfolio of CT as an outcome measure.

**Correlation with other surrogate endpoints**

In addition to correlations with true endpoints, CT parameters have also been linked to the following frequently used surrogate endpoints.  

**Spirometry parameters:** Several studies have shown a correlation between CT and spirometry parameters such as FEV₁, currently the most extensively validated surrogate endpoint in CF. FEV₁ is indirectly related to structural lung
damage, and has been shown to correlate with survival. (11, 78-88) From these studies we learned that CT scores are more sensitive than PFTs in detecting and monitoring onset and progression of CF lung disease.(51-52) Comparing the validation status of FEV₁ to CT, it can be concluded that most of the validation steps for CT have been addressed (Table 1).

*Lung clearance index (LCI):* CT scores have also been shown to correlate with LCI, a promising early marker derived from multiple breath washout.(89-91) The LCI has been shown to be more sensitive to detect lung disease than spirometry.(89, 92-95) In addition, LCI is reproducible (92, 95), and can be performed in infants.(96) Its narrow range in normal subjects makes it a suitable measure for long term follow up.(89, 92-95) A normal LCI has even been suggested to exclude structural changes on CT (89-90), a finding that was not supported by another study.(91) Its sensitivity to monitor progression of advanced lung disease however, has not yet been investigated. In addition, LCI has not been well validated against other true and surrogate endpoints.

*Inflammatory parameters:* The severity of structural changes on CT has been shown to correlate with inflammatory parameters in bronchoalveolar lavage (BAL). In a study in 17 children under the age of 4 years, regional distribution of airway disease was assessed using CT scans, and abnormalities were correlated to markers of lower airway inflammation provided by BAL. They found that in the lobe with greatest disease as indicated by CT, inflammatory markers were higher than in lobes with least disease.(58) These findings have been confirmed in two
other studies in infants with CF identified by newborn screening who underwent
BAL and CT scanning as part of an early surveillance program.(47, 50)
Thus, CT has been shown to be clearly linked to a number of clinically
meaningful outcome measures. These correlations are essential in the validation
of CT as surrogate endpoint in CF.

**Advantages of using CT in clinical trials**

In the previous paragraphs, we have shown that CT meets many of the
requirements for surrogate endpoints. CT is able to detect structural changes
relevant for CF patients, abnormalities on CT can be reproducibly quantified, and
standard CT scanning procedures have been defined. CT scores can improve
with treatment, although this requirement has not been completely met for
bronchiectasis. In addition, CT parameters have been linked to other clinically
meaningful outcome measures. Thus, CT can be considered quite validated as a
surrogate endpoint for CF-related lung disease. Using CT as surrogate endpoint
has the following advantages. First, CT can easily be performed in most CF
centers, as virtually all centers are equipped with a CT scanner. Second, CT can
be performed across all age ranges, including infants. Third, CT is the most
sensitive tool to detect early and regional disease. Compared with spirometry
parameters, CT is more sensitive to detect and monitor disease progression.(51-
52, 97) It has been estimated that the better sensitivity of CT relative to PFTs can
reduce sample size in clinical studies substantially.(5, 70) This would increase
the feasibility to run clinical trials in CF.
Disadvantages of using CT in clinical trials

Clearly, the use of chest CT in clinical trials has a number of disadvantages such as ionizing radiation, the need for general anesthesia in young children, costs, and the relative complicated procedures needed for the pressure-controlled volume scanning.

**Radiation:** CT exposes patients to ionizing radiation, which increases one’s natural life-long risk of cancer. Children are particularly at risk, as tissue and organs are growing and developing and therefore more sensitive to radiation. In addition, children have a longer life expectancy and therefore more time to manifest the oncogenic effects. At high exposures, the risk of cancer increases linearly with increasing dose. The relationship between radiation exposure and cancer risk from low-dose radiation (such as CT scanning) is less clear. In a study using a computational model, bi-annual chest CTs (mean dose 1 milli Sievert (mSv)) were shown to carry a low risk of radiation-induced mortality, with reported cumulative cancer mortalities of 1% at age 40 and 6% at age 65.\(^{(98)}\) In addition, the association between radiation dose from CT scans and leukaemia and brain tumors has been reported recently, with risk estimates in a similar range as those stated above.\(^{(99)}\) However, with increasing overall survival, life-long radiation exposure also increases, and thus the risk of radiation-induced cancer and mortality can become more meaningful. Therefore, protocols should aim to limit radiation to the absolute minimum needed to acquire images of sufficient quality. Six slice protocols with a mean dose of 0.19 mSv have been
described, and were found to be appropriate for evaluating bronchiectasis in pediatric patients.(100) However, limited slice CT reduces the ability to identify specific areas and their interval change, thus decreasing its ability to detect a therapeutic effect. Volumetric CT scanning will improve the ability to identify and compare specific structures on serial CTs. Currently, we can acquire volumetric CTs (inspiratory plus expiratory images) of sufficient quality for assessment of structural changes with a mean total effective dose near 1 mSv.(64) This is comparable to 1/3 of the annual US background radiation.(101) These doses can likely be further reduced in the near future.(64) Keeping the risk-benefit ratio of clinical trials in mind, CT should be considered as surrogate endpoint in studies aiming to slow the progression of CF lung disease.

Sedation/anesthesia: A second disadvantage is that CT scanning in infants and young children requires sedation (i.e., with chloral hydrate) or general anesthesia. These agents are very effective, with reported sedation failure rates of less than 1% using chloral hydrate.(102) In addition, they have a good safety profile. In two large retrospective studies reviewing records of children who underwent sedation for imaging studies, adverse events were reported in 0.85% and 0.42% of patients respectively.(102-103) Recent animal studies and experimental data have raised concerns about the potential neuro-toxicity of sedation and anesthetic agents and their effect on the developing brain.(104-105) To date however, there is still insufficient evidence to determine whether anesthetics are harmful to the developing human brain.(106) Thus, the risks of sedation are low, but not zero. To minimize these risks, patients should be
carefully selected and monitored (107), and sedated using the most suitable technique.(108) The advantage of sedation/anesthesia is that lung volumes can be standardized. Sedation/anesthesia can be avoided using recently developed ultra fast CT scanners. A comparative study showed that excellent image quality could be achieved in infants using a second generation (2 x 128 slices) dual source CT without sedation with a radiation dose similar to that of conventional scanners.(109) However, CTs acquired during free breathing are usually taken at a volume level near FRC. Whether the sensitivity and accuracy of these scans to detect structural abnormalities is comparable to that of volume controlled inspiratory and expiratory scans has not yet been investigated.

Costs: A third disadvantage of CT is the high costs. CT as surrogate endpoint will increase the costs of an intervention study. However, CT likely reduces the required sample size, which ultimately reduces the total costs of a trial. To the best of our knowledge, the costs and cost-effectiveness of CT in CF clinical trials has never been investigated. More research is needed to investigate this further.

Volume control: A fourth disadvantage is that CT protocols including PCV techniques or a spirometer are relatively complicated and requires a trained team. Children of 5 years and older need to be trained and coached by lung function technicians before and during scanning. In addition, a spirometer is required for use in the CT room. Procedures involving children below the age of 5 require personnel experienced in pediatric sedation/anesthesia and the PCV procedure. Furthermore, adequate time slots must be available to execute the protocol. This may limit its feasibility in multicenter trials.
Chest CT for clinical use

Over the last decade, more and more centers have adopted the routine to do a chest CT every other year at the time of the annual checkup in combination with PFTs. (51) A few centers perform a CT scan every 3rd year. (52) This strategy is considered useful to determine the long term efficacy of maintenance therapy. The most important reason for this routine is that chest CT is our current gold standard to detect bronchiectasis. (46) As discussed, chest CT provides a more sensitive and accurate estimate of the severity and progression of CF lung disease compared to PFTs. (51-52) In 50% of patients, discordance is observed between PFTs and chest CT data. Several studies showed that, despite stable spirometry parameters, CT scores show an annual progression ranging from 0.8% to 2.2% for total CT score, 1% to 1.7% for bronchiectasis, 1% to 1.4% airway for wall thickening, 0.4% to 0.65% for mucus plugging, and 1% to 7% for trapped air. (51-52, 74-75)

Clearly, a multitude of modalities including PFTs are used in addition to monitor CF lung disease, and each of these modalities contribute to clinical decision making. For this reason, it is nearly impossible to determine the impact of chest CT to the prognosis contribution. A question often raised is in how far CT affects clinical management. In case progression of bronchiectasis on CT is observed, adjustment to the treatment for the 2 years following the CT study should be made with the aim to stop further progression. For clinical management the CT interval of 2 years is in general too long for the timely identification of fast progressing patients. Due to the radiation related to chest CT, this
modality should not be routinely used to monitor CF lung disease over short time intervals like for the identification of exacerbations. To monitor the course of CF lung disease for the time interval between bi-annual chest CTs, lung function measurements like spirometry and LCI will continue to play an important role. When a more intensive monitoring protocol of lung structure is needed for a high risk patient, one can consider to alternate routine bi-annual chest-CT with a bi-annual chest MRI or with a ultra-low dose expiratory CT scan to reduce cumulative radiation dose. Further reductions in radiation dose in the near future will increase our options.

**Future research**

In this review, we have presented a large body of evidence supporting the use of chest CT as surrogate endpoint in clinical studies. The follow further improvements can be made to increase the accuracy and sensitivity of the technique.

*Calibration, standardization:* CT has already shown its feasibility in multi-center trials. These trials demonstrated the importance of defining a detailed protocol to improve compliance, and to document the ability of the centers to apply this protocol. (111) CT studies require well defined standard operating procedures to guarantee optimal image resolution and the use of the correct reconstruction algorithms. Since CT technology is moving fast, the procedures will require frequent updating.
**Image analysis:** Standardization is also important for CT quantification. Agreement on the most appropriate scoring systems for the different CF cohorts (infants, children and adults) is necessary, in addition to commercial availability. Currently available fully automated systems require validation in CF, and their usefulness in longitudinal studies should be determined. Such systems can further improve the sensitivity and accuracy of chest CT as a surrogate end point.

**Further validation steps:** An important next step will be to study the efficacy of antibiotics, mucolytics or disease modifiers on the development of bronchiectasis to further validate the response to treatment. In addition, it will be necessary to demonstrate that the effect of an intervention on the CT score predicts the effect on true clinical endpoints such as RTE-R and QoL. This validation is specific to the class of intervention, and the CT score studied. Different CT scores may capture different types of structural changes, and therefore may affect different primary endpoints. Furthermore, it is important to establish the minimal clinically relevant changes for each CT score. Thus, more comparative studies are needed. More longitudinal studies will help to further establish the trends of the different CT features over time. This will help defining the required time span for intervention studies. Ideally, a profile should emerge defining the therapeutic indications and target populations where CT will be most useful and likely to show change.
Summary

In this review, we showed that CT scores meet all critical requirements for surrogate endpoints. CT is able to detect structural changes highly relevant for CF patients using standardized image acquisition protocols, and the abnormalities detected using CT can be reproducibly quantified. Most CT scores can improve with treatment, although this needs to be further established for the bronchiectasis score. In addition, CT has been linked to a number of clinically meaningful outcome measures. CT can be performed across all age ranges, is the most sensitive tool to assess mild and regional disease, and likely reduces sample size requirements for clinical studies compared to other recognized endpoints. The disadvantage of chest CT in clinical studies however, include the radiation exposure to the patient, the need for sedation in infants, high costs, and the more complicated procedures when volume control is needed. Further innovations in image analysis techniques will help to further improve sensitivity and accuracy of chest CT to monitor CF lung disease. Remaining challenges in the use of CT as surrogate endpoint in clinical trials comprise standardization of CT protocols, procedures and quantification, better assessment of accuracy, and reliability of CT scores, better understanding of the association of CT scores with clinical outcomes, together with assessment of its feasibility in multi-center settings.
### Table 1. Overview of the literature published on the validation of the surrogate endpoints $FEV_1$ and CT.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$FEV_1$ References</th>
<th>CT References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence and severity of CF lung disease</td>
<td>(18), (12), (123), (113),</td>
<td>(46), (29), (49), (47),</td>
</tr>
<tr>
<td></td>
<td>(14), (115), (116),</td>
<td>(50), (120), (52),</td>
</tr>
<tr>
<td></td>
<td>(19), (86), (117),</td>
<td>(114), (97), (120)</td>
</tr>
<tr>
<td></td>
<td>(118), (83), (119)</td>
<td></td>
</tr>
<tr>
<td>Response to treatment in CF</td>
<td>(121), (122), (123),</td>
<td>(55), (56), (57), (58)</td>
</tr>
<tr>
<td></td>
<td>(56), (35), (124),</td>
<td>(59), (60), (35), (62)</td>
</tr>
<tr>
<td></td>
<td>(125), (126), (127),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(128), (129), (130),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(131), (132), (133),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(134), (135), (136),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(137), (138), (139),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(140), (141), (142)</td>
<td></td>
</tr>
<tr>
<td>Reproducibility of the measurement</td>
<td>(143), (144), (145),</td>
<td>(28), (29)</td>
</tr>
<tr>
<td></td>
<td>(146), (147), (148),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(149), (150), (151)</td>
<td></td>
</tr>
<tr>
<td>Link to true outcomes in CF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory tract exacerbations</td>
<td>(75), (74), (10), (84),</td>
<td>(75), (74)</td>
</tr>
<tr>
<td></td>
<td>(121), (152), (115)</td>
<td></td>
</tr>
<tr>
<td>Survival</td>
<td>(11), (78), (79), (80),</td>
<td>(77)</td>
</tr>
<tr>
<td></td>
<td>(81), (82), (83), (84),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(85), (86), (87), (88)</td>
<td></td>
</tr>
<tr>
<td>Quality of life</td>
<td>(153), (154), (155),</td>
<td>(166), (76)</td>
</tr>
<tr>
<td></td>
<td>(156), (157), (158),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(159), (160), (161),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(162), (163), (164),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(165)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Image showing the dissociation between lung function and lung structure assessed by computed tomography (CT). Shown is a slice of a routine CT scan of a cystic fibrosis patient, performed during annual check up when clinically stable. This slice clearly shows bronchiectasis and bronchial wall thickening (white arrows), and mucus plugging (black arrow). Spirometry parameters at the time of the CT however, were in the normal ranges, with a forced expiratory volume in 1 second of 98%-predicted.
References

1. Cystic Fibrosis Foundation 2010 annual report.


34. Loeve, M., M. de Bruijne, V. Gorbunova, W.C. Hop, and H. Tiddens, Reversibility of trapped air on CT scans of CF patients, an automated approach. Abstract#201, in European Cystic Fibrosis Society Conference. 2011: Hamburg, Germany.


are Associated with Future Lung Disease Progression in Children with CF. Am J 
Respir Crit Care Med, 2011.

Tiddens, J. Aldrich, H.O. Coxson, and D.D. Sin, Estimation of cancer mortality 
associated with repetitive computed tomography scanning. Am J Respir Crit Care 

99. Pearce, M.S., J.A. Salotti, M.P. Little, K. McHugh, C. Lee, K.P. Kim, N.L. Howe, 
C.M. Ronckers, P. Rajaraman, A.W. Craft, L. Parker, and A. Berrington de 
Gonzalez, Radiation exposure from CT scans in childhood and subsequent risk of 

100. O'Connor, O.J., M. Vandeleur, A.M. McGarrigle, N. Moore, S.R. McWilliams, 
S.E. McSweeney, M. O'Neill, M. Ni Chroinin, and M.M. Maher, Development of 
low-dose protocols for thin-section CT assessment of cystic fibrosis in pediatric 


102. Egelhoff, J.C., W.S. Ball, Jr., B.L. Koch, and T.D. Parks, Safety and efficacy of 
sedation in children using a structured sedation program. AJR Am J Roentgenol, 

Connor, and K.P. Mason, Adverse cardiovascular and respiratory events during 
p. 288-94.

104. Jevtovic-Todorovic, V., Developmental synaptogenesis and general anesthesia: a 

105. McCann, M.E. and S.G. Soriano, General anesthetics in pediatric anesthesia: 

106. Vutskits, L., Anesthetic-related neurotoxicity and the developing brain: shall we 

Wilson, and S. Work Group on, Guidelines for monitoring and management of 
pediatric patients during and after sedation for diagnostic and therapeutic 

Sedation for diagnostic and therapeutic procedures in children and young people: 

Achenbach, M. Uder, and T. Radkow, High-pitch spiral computed tomography: 
effect on image quality and radiation dose in pediatric chest computed 

Milla, A.S. Brody, J.P. Clancy, B. Ramsey, N. Hamblett, and A.E. Heald, 
Repeated adeno-associated virus serotype 2 aerosol-mediated cystic fibrosis


