Title

PRAGMA-CF: a quantitative structural lung disease CT outcome in young children with cystic fibrosis.

Authors

Tim Rosenow^{1,2}, Merel C.J. Oudraad^{3,4}, Conor P. Murray⁵, Lidija Turkovic², Wieying Kuo^{3,6}, Marleen de Bruijne^{7,8}, Sarath C. Ranganathan^{9,10,11}, Harm A.W.M. Tiddens^{3,6}, Stephen M. Stick^{2,12} on behalf of AREST CF.

Affiliations:

- 1. School of Paediatrics and Child Health, The University of Western Australia, Perth, Australia*
- 2. Telethon Kids Institute, The University of Western Australia, Perth, Australia*
- 3. Department of Pediatric Pulmonology, Erasmus MC-Sophia Children's Hospital, Rotterdam, The Netherlands*
- 4. Faculty of Medicine, University of Utrecht, Utrecht, The Netherlands*
- 5. Department of Diagnostic Imaging, Princess Margaret Hospital for Children, Western Australia, Australia*
- 6. Department of Radiology, Erasmus MC, Rotterdam, The Netherlands*
- 7. Biomedical Imaging Group Rotterdam, Departments of Radiology and Medical Informatics, Erasmus MC, Rotterdam, The Netherlands*
- 8. Department of Computer Science, University of Copenhagen, Denmark*
- 9. Infection and Immunity Theme, Murdoch Children's Research Institute, Melbourne, Australia*
- 10. Department of Respiratory Medicine, Royal Children's Hospital Melbourne, Australia*
- 11. Department of Paediatrics, University of Melbourne, Melbourne, Australia.*
- 12. Department of Respiratory and Sleep Medicine, Princess Margaret Hospital for Children, Perth, Australia*

Running Head

PRAGMA-CF: a quantitative CT outcome in CF

Corresponding author

Tim Rosenow

Ph: +61 8 9489 7818 | Fax: +61 8 9489 7700

Email: Tim.Rosenow@telethonkids.org.au

Contributions

Study conception and design: all authors equally.

Literature search: TR, HT.

CT analysis: TR, CM, MO.

Data/statistical analysis: TR, LT, SS.

Data interpretation: all authors equally.

Generation of figures: TR, LT.

Manuscript drafting and revision for important intellectual content, all authors equally.

Manuscript editing: all authors equally.

Accountability for all aspects of the work: all authors equally.

Final approval of the submitted manuscript: all authors.

Sources of income

Cystic Fibrosis Foundation (USA) (Project grant support), National Health and Medical Research Council (Australia) (Project grant support and Fellowships), and Cystic Fibrosis Australia (Project grant support).

AJRCCM subject category code

9.17 Pediatrics - Cystic Fibrosis: Translational & Clinical Studies

Manuscript word count

3071 / 3500

At a Glance Commentary

Scientific Knowledge on the Subject

Currently there are no standardised outcome measures appropriate for very young children with cystic fibrosis (CF). This effectively excludes infants and young children from clinical trials, in an era where new, potentially disease modifying drugs are becoming available. Chest computed tomography (CT) is the gold standard for detecting early CF lung disease, so a validated outcome measure from CT is urgently needed.

What This Study Adds to the Field

This study introduces PRAGMA-CF, a new quantitative outcome measure for chest CT appropriate for young children with early CF lung disease. It shows that PRAGMA-CF is repeatable, sensitive to early changes, and feasible for use as a clinical trial outcome.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

Declaration of interests:

Dr. Tiddens reports grants from CFF, during the conduct of the study; other from Roche, other from Pharmaxis, other from Novartis, grants from CFF, grants and other from Vertex, grants and other from Gilead, outside the submitted work; In addition, Dr. Tiddens has a patent ActivAero licensed and is heading the ErasmusMC-Sophia Children's Hospital core laboratory LungAnalysis.

Dr. de Bruijne reports grants from Netherlands Organisation for Scientific Research (NWO), during the conduct of the study; grants from Astra Zeneca, grants from Vertex, outside the submitted work; In addition, Dr. de Bruijne has a patent No US8811724 Classification of medical diagnostic images issued.

No other authors have any conflicts of interest to disclose.

ABSTRACT

Rationale: Chest computed tomography (CT) is the gold standard for demonstrating cystic fibrosis (CF) airways disease. However, there are no standardised outcome measures appropriate for children under 6 years.

Objectives: We developed the Perth-Rotterdam Annotated Grid Morphometric Analysis for CF (PRAGMA-CF), a quantitative measure of airways disease, and compared it to the commonly used CF-CT scoring method.

Methods: CT scans from the Australian Respiratory Early Surveillance Team for CF (AREST CF) cohort in Western Australia were included. PRAGMA-CF was performed by annotating a grid overlaid on ten axial slices for the presence of bronchiectasis, mucous plugging or other airway abnormalities (inspiratory scans) and trapped air (expiratory scans). The separate proportions of total disease (%Dis), bronchiectasis (%Bx) and trapped air (%TA) were determined. Thirty scans were used for observer reliability, and thirty paired scans obtained at 1 and 3-years old were used for comparison with a validated standard and biological plausibility.

Measurements and Main Results: Intraobserver, intraclass correlation coefficients (95% confidence interval) for %Dis, %Bx and %TA were 0.93 (0.86 – 0.97), 0.93 (0.85 – 0.96) and 0.96 (0.91 – 0.98), respectively. The change in %Dis ($p = 0.004$) and %Bx ($p = 0.001$) with PRAGMA-CF was related to neutrophil elastase (NE) presence at age 3, whereas only the change in bronchiectasis score was related to NE ($p < 0.001$) with CF-CT. Sample size calculations for various effect sizes are presented.

Conclusions: PRAGMA-CF is a sensitive and reproducible outcome measure for assessing the extent of lung disease in very young children with CF.

Word count: 249/250. Key words: Imaging, Medical;Pediatrics;Clinical Trials as Topic

INTRODUCTION

Structural lung disease (SLD) in cystic fibrosis (CF) begins early in life, is progressive, and is often the only evidence of respiratory disease in children less than 6 years of age.(1-4) Chest computed tomography (CT) is the gold standard for demonstrating CF-related SLD. However, there are no validated quantitative outcomes based on CT that are appropriate for young children with early lung disease. Therefore, in an era that promises new disease modifying therapies, an accurate and sensitive assessment tool is urgently needed both for monitoring disease progress and for clinical trials.

Several CT scoring systems for CF have been developed for adults and children over the age of 6, with CF-CT being the most comprehensively examined.(5-8) These methodologies are semi-quantitative, and have not been designed for assessing subtle appearances and low extents of structural changes found early in life.(3) In particular, the extent scores for components of SLD are generally based on the approximate proportion of the lung affected, with the smallest increment being one third to one half of a lobe.(3, 5) The majority of young children have isolated disease, only the mildest severity of lung abnormalities, and lobar disease extent that is well below 50%.(9-11) This creates, using currently available scoring systems, a binary outcome measure for each lobe in the majority of scans. Therefore, a quantitative measure that is sensitive to early SLD is essential for use in clinical trials or longitudinal assessment in this young age group.

This paper introduces the Perth-Rotterdam Annotated Grid Morphometric Analysis method (PRAGMA-CF): a novel quantitative measure of CT SLD that specifically targets early CF lung disease. We assessed the intra- and interobserver agreement for PRAGMA-CF, relations with factors known to be associated with the development and progression of SLD and compared the strengths of the associations using both PRAGMA-CF and CF-CT methods.

We hypothesised that during the first 3 years of life, the PRAGMA-CF approach compared to the CF-CT algorithm will have better repeatability, be more sensitive to disease progression and more strongly correlate with airway inflammatory biomarkers.

Some of the results of these studies have been previously reported in the form of abstracts.(12-14)

METHODS

Study population and image selection: The children in this study were participants in the Australian Respiratory Early Surveillance Team for Cystic Fibrosis (AREST CF) early surveillance programme. In order to determine the minimum number of annotated slices required for accurate analysis, we selected twenty volumetric inspiratory and expiratory scans representing a range of SLD extents (set_{slice}): ten scans from patients aged below 6 years with the highest CF-CT bronchiectasis score ("severe" cases), and ten scans randomly selected from patients aged 3 and under ("mild" cases). To determine intra- and interobserver repeatability, thirty additional volumetric

inspiratory and expiratory scans were randomly selected from patients aged below 6 years (set_{rep}). Random scan selection was performed by assigning all eligible CTs scan a pseudorandom number in the interval $[-1, 1]$, sorting the list of scans based on this value, and selecting the required number of scans with the lowest rank (ensuring that the same patient was no included more than once).

In order to assess biological plausibility and assess PRAGMA-CF as a potential clinical trial outcome for infants, we examined data from all patients who had paired volumetric inspiratory CT scans (and either volumetric or limited slice expiratory scans) acquired at both the ages of 1 and 3 years (set_{bio}).

CT image acquisition: Chest CT scans were obtained under general anaesthesia as previously described.(3)

CT scoring: CT scans were scored using a simplified version of the CF-CT scoring system(9, 15) and the PRAGMA-CF algorithm.

Quantitative CT analysis: We developed PRAGMA-CF based on the severe advanced lung disease method described by Loeve et al.(16) Using the thinnest slice reconstruction (0.8-1.0mm) of the inspiratory scan, ten slices equally spaced between the apex and the base of the lung were selected for annotation and overlaid with a square grid. We corrected for lung size by setting the grid cell size equal to $1/20^{\text{th}}$ of the lung width at the carina (rounded to the nearest mm), measured with electronic callipers using Myrian software (Intrasense, Montpellier, France). This size was chosen as it approximately corresponds to the size of the largest airways.

For each grid cell that was at least 50% covered by the lung field, any presence of airways disease was annotated according to a hierarchical system, from highest to lowest priority: bronchiectasis (outer edge bronchus-artery cross-sectional area ratio > 1); mucous plugging (high density airway occlusion or tree-in-bud appearance); bronchial wall thickening (assessed subjectively as airway walls that are thicker or have increased signal intensity relative to normal airways); atelectasis; or normal lung structure. Although only select slices were annotated, the observer was permitted to scroll through the entire lung volume to assist in classification, for example to distinguish between an occluded airway and an artery. Cells containing atelectasis were annotated but excluded from all analysis as they are likely related to general anaesthesia rather than pathology.⁽¹⁷⁾ The primary outcomes from this method are the volume proportion of the lung with airways disease (%Dis), determined by dividing the number of cells annotated with either bronchiectasis, mucous plugging or airway abnormality by the total number of annotated cells (excluding atelectasis), and the volume proportion of the lung with bronchiectasis (%Bx), determined by dividing the number of cells annotated with bronchiectasis by the total number of annotated cells (excluding atelectasis).

Trapped air was assessed with a similar methodology on the expiratory scans. A Minimum Intensity Projection (MinIP) reconstruction was performed prior to analysis with a slice thickness of 4 to 5 mm, depending on the clinical protocol at the time. This projection increases visibility of low intensity regions and facilitates visualisation of trapped air.^(18, 19) Cells were annotated according to whether trapped air represented 50% or more of the lung field in

the cell (trapped air) or less than 50% (healthy). The primary outcome was the volume proportion of the lung with trapped air (%TA), determined by dividing the number of trapped air cells by the total number of annotated cells.

Prior to 2010, clinical procedure was to acquire expiratory scans using a three-slice protocol. For these scans, the lung width was measured on the second (middle) slice, and all three slices were annotated for the presence of trapped air. Because MinIP reconstructions are not possible on three-slice scans, standard 1mm slice thickness lung reconstructions were used for all expiratory scans in set_{bio} to minimise the effect of scan technique on longitudinal assessment of trapped air.

Slice annotation interval: Scans from $\text{set}_{\text{slice}}$ were assessed using PRAGMA-CF, with slices annotated at 5mm intervals in order to determine the most accurate estimate of %Dis, %Bx and %TA. Every second annotated slice was then removed from analysis, representing PRAGMA-CF at 10mm intervals, and then again representing PRAGMA-CF at 20mm intervals. The maximum distance between annotated slices was determined by finding the largest interval that did not result in a significant change in PRAGMA-CF outcomes. Based on these results, a fixed number of 10 annotated slices was chosen as the standard for PRAGMA-CF. This standard was used for assessing all other data sets, and takes approximately 20 minutes to analyse.

Observer reliability: The primary observer for this study (TR) was a PhD student, trained in CF-CT scoring, with more than 3 years of experience in early CF lung disease. A second observer (MO), a medical student previously trained in CF-CT scoring, was provided with a detailed instruction manual on PRAGMA-CF and a practise set of 6 batches of 5 CT scans to annotate.

Detailed feedback based on visual assessment of the annotations was given after each batch.

To determine interobserver reliability, scans from set_{rep}, were de-identified and scored in random order with PRAGMA-CF independently by both observers.

This set was re-randomised and rescored after one month (to reduce recall bias) by the primary observer in order to determine intraobserver repeatability.

Biological validation: We reasoned that any new assessment of SLD should reflect known pathobiological factors. We therefore compared relations between inflammatory biomarkers and both the PRAGMA-CF and CT-CT outcomes. Following the CT scan, bronchoalveolar lavage (BAL) was performed to determine interleukin-8 (IL8) concentration, neutrophil count and neutrophil elastase (NE) presence, as previously described.(2, 20)

Statistical analysis: Slice annotation interval scores were analysed pairwise by spacing, using the intraclass correlation coefficient (ICC). Inter- and intraobserver reliability was assessed using the ICC (two-way mixed effects model) and Bland-Altman analysis.(21) In general, ICC between 0.4 and 0.6 is in general considered moderate, between 0.6 and 0.8 good, and greater than 0.8 excellent.(22) CF-CT scores were treated as continuous variables for all analyses. PRAGMA-CF and CF-CT scores were compared against binary and continuous outcomes using the Wilcoxon rank-sum test and Spearman's rank correlation (ρ_s), respectively.

All analyses were performed using Stata version 13.0 (Stata Corp., College Station, TX).

RESULTS

Slice annotation interval: Median (range) PRAGMA-CF outcomes for $\text{set}_{\text{slice}}$ for %Dis, %Bx and %TA at 5mm intervals were 3.54 (0.00 – 23.76), 1.76 (0.00 – 16.91) and 2.05 (0.00 – 51.51), respectively. ICCs between 5mm and 10mm, and 5mm and 20mm intervals exceeded 0.99 for all outcome measures (see online supplement). Ten slices was chosen as the standard for PRAGMA-CF, so that the volume of lung annotated is standardised across lung heights, and thus ensuring that annotation intervals are < 20mm for this age group.

Observer reliability: Median (range) PRAGMA-CF outcomes for set_{rep} for %Dis, %Bx and %TA were 3.99 (0.72 – 11.59), 0.00 (0.00 – 6.84) and 1.52 (0.00 – 55.36), respectively. The ICC was greater than 0.8 for all outcomes, both for intra- and interobserver reliability (Table 1).

Biological validation: Thirty patients had paired volumetric inspiratory scans at the ages of 1 and 3 (Table 2). Of these, six patients had volumetric expiratory scans at age 1, and 25 at age 3; the remainder of expiratory scans were three-slice. A change in expiratory scan methodology (volumetric vs three-slice) between visits was not related to the change in %TA ($p = 0.576$).

Inflammation: There were significant cross-sectional relationships between all *inspiratory* CT outcomes and inflammatory markers (Table 3). For *expiratory* scans, %TA was significantly related to neutrophil count and NE presence; whereas no relationships were found with CF-CT trapped air score. Longitudinal change in %Dis was significantly related to neutrophil count and

NE presence at age 3; however total CF-CT score was related only to IL-8 concentration. The change in both %Bx and CF-CT bronchiectasis score were significantly related to all inflammatory markers at age 3 (Table 4). Additional graphical data are provided in the online supplement.

Bronchiectasis progression: %Dis at age 1 was significantly related to the change in %Bx ($\rho_s = 0.44$, $p = 0.016$). This relationship was also present between total CF-CT score at age 1 and the change in CF-CT bronchiectasis score ($\rho_s = 0.36$, $p = 0.050$). No other CT outcomes at age 1 were related to bronchiectasis progression.

Trapped air: There were significant correlations between %TA and %Dis ($\rho_s = 0.50$, $p < 0.001$), and %TA and %Bx ($\rho_s = 0.42$, $p < 0.001$). These relationships were not significant between CF-CT trapped air score and total score ($\rho_s = 0.22$, $p = 0.091$) or bronchiectasis score ($\rho_s = 0.17$, $p = 0.197$). Longitudinally, the change in %TA was significantly related to the change in %Dis ($\rho_s = 0.62$, $p < 0.001$) and the change in %Bx ($\rho_s = 0.47$, $p = 0.008$). These relationships were not significant between changes in CF-CT trapped air score and total score ($\rho_s = 0.12$, $p = 0.533$) or bronchiectasis score ($\rho_s = 0.30$, $p = 0.107$).

Sample size estimates: We performed power calculations based on a range of effect sizes. The mean (SD) %Dis at 3 years was 2.62 (2.59); we calculated sample sizes based on trials aiming to reduce this by 75, 50 and 25%. In addition, 46% of patients either developed or had progression in the extent of

bronchiectasis; we also calculated sample sizes aimed at reducing this proportion to 10, 20 and 30% (Table 5).

DISCUSSION

This study presents the first quantitative method for assessing SLD on chest CT designed specifically for young children and infants with CF. We confirmed that although the prevalence of SLD early in life is high,(2, 3, 23) the extent is relatively low, with a median proportion of lung volume affected being 0.87% and 1.86% at age 1 and 3, respectively. The most important findings of this study are that the quantitative PRAGMA-CF outcomes:

- Have high intra- and interobserver agreement;
- are better correlated to neutrophilic inflammation than CF-CT scores; and,
- show stronger relationships between structural changes and trapped air progression.

Measures of agreement: ICCs were all greater than 0.90 for intraobserver and greater than 0.85 for interobserver comparisons, similar to previously published intraobserver repeatability for simplified CF-CT scores of bronchiectasis and trapped air.(9) The mean differences between raters and between the first and second scoring attempts, although significant, were small in magnitude, and likely skewed by a few high-scoring patients. Although there is no direct comparison between CF-CT outcomes and %Dis, it is as repeatable as the other PRAGMA-CF outcomes. PRAGMA-CF therefore is at least as repeatable as CF-CT, but over a much more sensitive

range (several percent of the lung compared to a binary per-lobe score of present/absent).

Biological validation

Inflammation: The number of patients with detectable NE was low, and therefore further studies are required in order to determine the true relationship between CT and neutrophil elastase. Cross-sectional relationships between inflammatory markers were similar for PRAGMA-CF and CF-CT scores, with the exception of trapped air, where PRAGMA-CF was better correlated. Longitudinally, the change in %Dis was related to the presence of NE and the neutrophil count at age 3, whereas the change in total CF-CT score was only related to IL-8. Because CF-CT is not quantitative, it is not clear how much of the total score relationship is driven by the bronchiectasis component. The change in both %Bx and the CF-CT bronchiectasis score were significantly related to all inflammatory markers at age 3 (Table 4). This suggests that PRAGMA-CF is more sensitive than CF-CT derived scores in assessing early changes in lung structure due to inflammation.

Progression: %Dis at the age of 1 was significantly related to the change in %Bx over the two years, suggesting that patients with worse baseline disease have faster progression of bronchiectasis. Therefore, PRAGMA-CF performed at the age of 1 year can potentially be used to identify patients at high risk for disease progression.

Trapped air: Trapped air assessed using PRAGMA-CF was significantly related to other CT outcomes (both cross-sectionally and longitudinally), but not when assessed with CF-CT scores. Although the relationship between trapped air and *bronchiectasis* progression has been previously demonstrated in children,(1, 24) the progression of SLD in general and trapped air in children aged below 3 years has not been observed. Our data suggest that this is because of the semi-quantitative nature of CF-CT and related scoring systems, rather than pathobiological reasons.

Sample size estimates: These data suggest that PRAGMA-CF could be an appropriate clinical trial outcome in young children, particularly for therapies aimed at reducing early inflammatory SLD. Because clinical trials have not previously been available for infants, the clinical impact of SLD extent in infancy is unknown. However, the sample sizes presented for a range of potential effect sizes suggest that multicentre trials in children between 1 and 3 years run over one to two years are feasible, with PRAGMA-CF as an outcome.

Further development: Grid cell size was arbitrarily chosen to be 1/20th of the lung width. This size was selected as it approximately represents the size of the largest assessable airway in the lung. Although objective comparisons between grid resolutions could determine the optimum cell size, the current value is already sufficient to demonstrate the utility of PRAGMA-CF. Future studies should be undertaken to assess the repeatability of PRAGMA-CF in a variety of age groups and disease severities, as it is possible that patients

with mild disease or of young age, as in this study, may have a larger *relative* variability of PRAGMA-CF outcomes compared to older subjects with more severe disease. PRAGMA-CF lends itself to automations, such as by using textural analysis(25), which will mitigate observer variability. However, even without further development, significant relationships between %Dis at age 1 and inflammation and bronchiectasis at age 3 were present, suggesting that PRAGMA-CF results at 1 year are likely to be clinically relevant and therefore relevant as a trial endpoint.

A limitation of the study with regard to trapped air assessments was that some expiratory scans were performed using a three slice protocol. Because we found no statistical difference in %TA in patients who had the same expiratory scan method on both visits, as compared to those who had a different method, we pooled the data for biological validation. Another limitations is that that MinIP reconstructions were not used for biological validation. Because MinIP reconstructions improve visualisation of low intensity regions, and hence reduce observer variability, we are likely to have underestimated any relationships between disease markers and PRAGMA-CF trapped air measurements.

The CF-CT scoring system has been used to evaluate SLD in young children with CF,(3, 26-28) but is relatively insensitive to mild disease. Our data using PRAGMA-CF demonstrate that reliable quantitative estimates of lung disease can be obtained in young children. These observations are significant because young children may stand to benefit the most from new, disease

modifying drugs that are becoming available, where until now, there has been a lack of a suitable outcome measures to enable clinical trials.(15) Our data suggest that PRAGMA-CF is a promising candidate as a standardised outcome measure for use in intervention studies and clinical trials in young children.

We believe that the circumstances now exist to make CT a rational choice as a primary outcome measure in clinical trials involving very young children with CF. Technological advances have resulted in improvements in CT image quality for a given radiation dose and improvements are likely to continue.(29) In an intervention study that requires images at two time points, the estimated excess risk of cancer is so low as to be incalculable.(30-32) Furthermore, using PRAGMA-CF, the sample sizes required for an intervention study with a relatively short duration i.e., two years, and the ubiquitous access at CF centres to CT scanners, mean that multiple studies can be undertaken simultaneously around the world. This is an important consideration given the low incidence of CF and the exciting prospects that now exist with regard to the development of disease modifying interventions.

Acknowledgements: We thank the Department of Diagnostic Imaging at Princess Margaret Hospital for Children for assistance with the acquisition and collation of CT scans, and Marcel Koek from Erasmus MC for the development of the grid software used for CT annotation.

REFERENCES

1. Mott LS, Park J, Murray CP, Gangell CL, de Klerk NH, Robinson PJ, Robertson CF, Ranganathan SC, Sly PD, Stick SM, Arest CF. Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* 2012; 67: 509-516.
2. Sly PD, Brennan S, Gangell C, de Klerk N, Murray C, Mott L, Stick SM, Robinson PJ, Robertson CF, Ranganathan SC, Australian Respiratory Early Surveillance Team for Cystic F. Lung disease at diagnosis in infants with cystic fibrosis detected by newborn screening. *American journal of respiratory and critical care medicine* 2009; 180: 146-152.
3. Stick SM, Brennan S, Murray C, Douglas T, von Ungern-Sternberg BS, Garratt LW, Gangell CL, De Klerk N, Linnane B, Ranganathan S, Robinson P, Robertson C, Sly PD, Australian Respiratory Early Surveillance Team for Cystic F. Bronchiectasis in infants and preschool children diagnosed with cystic fibrosis after newborn screening. *The Journal of pediatrics* 2009; 155: 623-628 e621.
4. Sly PD, Gangell CL, Chen L, Ware RS, Ranganathan S, Mott LS, Murray CP, Stick SM, Investigators AC. Risk factors for bronchiectasis in children with cystic fibrosis. *N Engl J Med* 2013; 368: 1963-1970.
5. Brody AS, Klein JS, Molina PL, Quan J, Bean JA, Wilmott RW. High-resolution computed tomography in young patients with cystic fibrosis: distribution of abnormalities and correlation with pulmonary function tests. *The Journal of pediatrics* 2004; 145: 32-38.
6. de Jong PA, Ottink MD, Robben SG, Lequin MH, Hop WC, Hendriks JJ, Pare PD, Tiddens HA. Pulmonary disease assessment in cystic fibrosis: comparison of CT scoring systems and value of bronchial and arterial dimension measurements. *Radiology* 2004; 231: 434-439.
7. Loeve M, Gerbrands K, Hop WC, Rosenfeld M, Hartmann IC, Tiddens HA. Bronchiectasis and pulmonary exacerbations in children and young adults with cystic fibrosis. *Chest* 2011; 140: 178-185.
8. Robinson TE, Leung AN, Chen X, Moss RB, Emond MJ. Cystic fibrosis HRCT scores correlate strongly with Pseudomonas infection. *Pediatric pulmonology* 2009; 44: 1107-1117.
9. Mott LS, Park J, Gangell CL, de Klerk NH, Sly PD, Murray CP, Stick SM, Australian Respiratory Early Surveillance Team for Cystic Fibrosis Study G. Distribution of early structural lung changes due to cystic fibrosis detected with chest computed tomography. *The Journal of pediatrics* 2013; 163: 243-248 e241-243.
10. Wainwright CE, Vidmar S, Armstrong DS, Byrnes CA, Carlin JB, Cheney J, Cooper PJ, Grimwood K, Moodie M, Robertson CF, Tiddens HA, Investigators AS. Effect of bronchoalveolar lavage-directed therapy on Pseudomonas aeruginosa infection and structural lung injury in children with cystic fibrosis: a randomized trial. *JAMA : the journal of the American Medical Association* 2011; 306: 163-171.
11. Thia LP, Calder A, Stocks J, Bush A, Owens CM, Wallis C, Young C, Sullivan Y, Wade A, McEwan A, Brody AS, London Cystic Fibrosis C. Is chest CT useful

- in newborn screened infants with cystic fibrosis at 1 year of age? *Thorax* 2014; 69: 320-327.
12. Rosenow T, Tiddens H, de Bruijne M, Berry L, Stick S. Quantitation of chest CT abnormalities in early life CF: back to the basics. The 27th Annual North American Cystic Fibrosis Conference. Salt Lake City, Utah, USA: Wiley-Blackwell 111 River St, Hoboken 07030-5774, NJ USA; 2013. p. 340-340.
 13. Rosenow T, Tiddens H, De Bruijne M, Stick S. PRAGMA: A new method of quantifying structural lung disease in young children with cystic fibrosis. Thoracic Society of Australia and New Zealand Annual Scientific Meeting: Wiley-Blackwell 111 River St, Hoboken 07030-5774, NJ USA; 2014. p. 48-48.
 14. Rosenow T, Tiddens H, Oudraad M, de Bruijne M, Stick S. PRAGMA: Further support for use as a quantitative CT outcome measure. The 28th Annual North American Cystic Fibrosis Conference. Atlanta, Georgia, USA: Wiley-Blackwell 111 River St, Hoboken 07030-5774, NJ USA; 2014. p. 295-295.
 15. Stick SM. Clinical trials in infants with cystic fibrosis detected following newborn screening. *Paediatric respiratory reviews* 2008; 9: 176-180.
 16. Loeve M, van Hal PT, Robinson P, de Jong PA, Lequin MH, Hop WC, Williams TJ, Nossent GD, Tiddens HA. The spectrum of structural abnormalities on CT scans from patients with CF with severe advanced lung disease. *Thorax* 2009; 64: 876-882.
 17. Newman B, Krane EJ, Gawande R, Holmes TH, Robinson TE. Chest CT in children: anesthesia and atelectasis. *Pediatric radiology* 2014; 44: 164-172.
 18. Bhalla M, Naidich DP, McGuinness G, Gruden JF, Leitman BS, McCauley DI. Diffuse lung disease: assessment with helical CT--preliminary observations of the role of maximum and minimum intensity projection images. *Radiology* 1996; 200: 341-347.
 19. Fotheringham T, Chabat F, Hansell DM, Wells AU, Desai SR, Guckel C, Padley SP, Gibson M, Yang GZ. A comparison of methods for enhancing the detection of areas of decreased attenuation on CT caused by airways disease. *Journal of computer assisted tomography* 1999; 23: 385-389.
 20. Fuchs SI, Gappa M, Eder J, Unsinn KM, Steinkamp G, Ellemunter H. Tracking Lung Clearance Index and chest CT in mild cystic fibrosis lung disease over a period of three years. *Respiratory medicine* 2014; 108: 865-874.
 21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-310.
 22. Loeve M, de Bruijne M, Hartmann IC, van Straten M, Hop WC, Tiddens HA. Three-section expiratory CT: insufficient for trapped air assessment in patients with cystic fibrosis? *Radiology* 2012; 262: 969-976.
 23. Mott LS, Park J, Murray CP, Gangell CL, de Klerk NH, Robinson PJ, Robertson CF, Ranganathan SC, Sly PD, Stick SM. Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* 2011.
 24. Tepper LA, Utens EM, Caudri D, Bos AC, Gonzalez-Graniel K, Duivenvoorden HJ, van der Wiel EC, Quittner AL, Tiddens HA. Impact of bronchiectasis and trapped air on quality of life and exacerbations in cystic fibrosis. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* 2013; 42: 371-379.
 25. Ciompi F, Palaoroutas A, Loeve M, Pujol O, Radeva P, Tiddens H, de Bruijne M. Lung tissue classification in severe advanced cystic fibrosis from CT scans. Fourth International Workshop on Pulmonary Image Analysis. p. 57-68.

26. Brody AS, Kosorok MR, Li Z, Broderick LS, Foster JL, Laxova A, Bandla H, Farrell PM. Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. *J Thorac Imaging* 2006; 21: 14-21.
27. Calder AD, Bush A, Brody AS, Owens CM. Scoring of chest CT in children with cystic fibrosis: state of the art. *Pediatric radiology* 2014; 44: 1496-1506.
28. Tiddens HA, Rosenow T. What did we learn from two decades of chest computed tomography in cystic fibrosis? *Pediatric radiology* 2014; 44: 1490-1495.
29. Newell JD, Jr., Fuld MK, Allmendinger T, Sieren JP, Chan KS, Guo J, Hoffman EA. Very Low-Dose (0.15 mGy) Chest CT Protocols Using the COPDGene 2 Test Object and a Third-Generation Dual-Source CT Scanner With Corresponding Third-Generation Iterative Reconstruction Software. *Investigative radiology* 2015; 50: 40-45.
30. Kuo W, Ciet P, Tiddens HA, Zhang W, Guillerman RP, van Straten M. Monitoring cystic fibrosis lung disease by computed tomography. Radiation risk in perspective. *American journal of respiratory and critical care medicine* 2014; 189: 1328-1336.
31. National Research Council . Committee to Assess Health Risks from Exposure to Low Levels of Ionizing Radiation. Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII-Phase 2. Washington, DC; 2005.
32. The American Association of Physicists in Medicine. AAPM Position Statement on Radiation Risks from Medical Imaging Procedures. 2011 12/13/2011 [cited 2014 18/11/2014].

TABLES

Intraobserver	ICC	Mean difference*
%Dis	0.929 (0.857 – 0.966)	n/s
%Bx	0.928 (0.854 – 0.965)	0.32 (p = 0.005)
%TA	0.960 (0.916 – 0.981)	2.03 (p = 0.001)
Interobserver		
%Dis	0.851 (0.710 – 0.926)	0.78 (p = 0.006)
%Bx	0.907 (0.815 – 0.955)	-0.38 (p = 0.006)
%TA	0.938 (0.873 – 0.970)	n/s

Table 1: Intraclass correlation (ICC) coefficients (95% confidence interval) and mean difference between observations (p-value from t-test) for intra- and inter-observer reliability of PRAGMA-CF outcomes in 30 CT scans.

* Mean difference between re-scores (intra) or raters (inter) and p-values. n/s = not significant

<i>Clinical Characteristics</i>	
Sex (M:F)	10:20
Severe genotype ^{#*}	22/27 (81%)
Pancreatic suff [#]	7 (23%)
Resp symptoms[#]	
- Age 1	6 (20%)
- Age 3	7 (25%)
<i>PRAGMA-CF</i>	
%Disease	
- Age 1	0.79 (0.00 – 4.88)
- Age 3	1.86 (0.00 – 10.53)
- Change	0.82 (-1.06 – 5.64)
%Bronchiectasis	
- Age 1	0.00 (0.00 – 0.62)
- Age 3	0.00 (0.00 – 6.72)
- Change	0.00 (-0.22 – 6.36)
%Trapped Air	
- Age 1	4.21 (0.00 – 29.79)
- Age 3	1.32 (0.00 – 59.79)
- Change	-1.63 (-20.29 – 47.34)
<i>CF-CT scores</i>	
Total	
- Age 1	4 (0 – 19)
- Age 3	11 (0 – 29)
- Change	7 (-1 – 16)
Bronchiectasis	
- Age 1	0 (0 – 5)
- Age 3	1 (0 – 10)
- Change	1 (0 – 8)
Trapped air	
- Age 1	2 (0 – 5)
- Age 3	1 (0 – 11)
- Change	0 (-5 – 10)
<i>Inflammation</i>	
Neutrophils	
- Age 1	1.06 (0.03 – 31.61)
- Age 3	4.03 (0.00 – 223)
IL-8	
- Age 1	0.48 (0.05 – 5.84)
- Age 3	1.38 (0.05 – 30.95)
NE present[#]	
- Age 1	2 / 30 (7%)
- Age 3	6 / 30 (20%)

Table 2: CT and inflammatory outcomes from set_{bio}, n=30, presented as median (range) unless otherwise specified. All PRAGMA-CF outcomes represent percentages of lung volumes. Total CF-CT score is out of 36, bronchiectasis and trapped air

subscores are each out of 12. Neutrophils = cell count / mL x 10^6 ; IL-8 = ng/mL; NE = neutrophil elastase presence; suff = sufficient. * = Severe genotype based on mutation class, data not available on 3 patients. # = data presented as prevalence (%).

	IL-8 concentration	Neutrophil count	NE presence
PRAGMA-CF			
%Disease			
- Age 1	0.39 (0.035)*	0.37 (0.045)*	0.025*
- Age 3	0.39 (0.036)*	0.44 (0.015)*	0.002*
- Combined	0.40 (0.002)*	0.38 (0.003)*	< 0.001*
%Bronchiectasis			
- Age 1	0.21 (0.264)	0.15 (0.426)	0.014*
- Age 3	0.48 (0.069)	0.66 (<0.001)*	<0.001*
- Combined	0.45 (<0.001)*	0.48 (0.001)*	< 0.001*
%Trapped Air			
- Age 1	0.09 (0.622)	0.38 (0.041)*	0.243
- Age 3	0.27 (0.142)	0.39 (0.035)*	0.032*
- Combined	0.12 (0.367)	0.36 (0.005)*	0.017*
CF-CT Scores			
Total			
- Age 1	0.38 (0.040)*	0.28 (0.137)	0.153
- Age 3	0.54 (<0.001)*	0.52 (0.003)*	0.002*
- Combined	0.55 (< 0.001)*	0.40 (0.002)*	<0.001*
Bronchiectasis			
- Age 1	0.20 (0.295)	0.15 (0.444)	0.124
- Age 3	0.41 (0.023)*	0.63 (<0.001)*	0.001*
- Combined	0.41 (0.001)*	0.43 (0.001)*	<0.001*
Trapped Air			
- Age 1	-0.03 (0.865)	-0.12 (0.518)	0.232
- Age 3	0.36 (0.052)	0.21 (0.260)	0.651
- Combined	0.18 (0.168)	0.06 (0.643)	0.938

Table 3: Cross-sectional relationships between CT outcomes and inflammatory markers in 60 scans from 30 patients. IL-8 and neutrophil data presented as Spearman's ρ_s (p-value). NE presence presented as p-values from Wilcoxon's rank-sum analysis. NE = Neutrophil elastase; * = $p < 0.05$.

	IL-8 concentration	Neutrophil count	NE presence
PRAGMA-CF			
Δ%Disease	0.18 (0.34)	0.41 (0.025)*	0.0043*
Δ%Bronchiectasis	0.41 (0.026)*	0.63 (0.0002)*	0.001*
Δ%Trapped Air	0.06 (0.74)	0.18 (0.33)	0.058
CF-CT Scores			
ΔTotal	0.37 (0.047)*	0.31 (0.096)	0.11
ΔBronchiectasis	0.43 (0.016)*	0.59 (0.0006)*	0.0008*
ΔTrapped Air	0.30 (0.10)	0.32 (0.09)	0.64

Table 4: Longitudinal relationships between the change (Δ) in CT outcomes between over two years and inflammatory markers at age 3 in 30 patients. IL-8 and neutrophil data presented as Spearman's ρ_s (p-value). NE presence presented as p-values from Wilcoxon's rank sum analysis. NE = Neutrophil elastase; * = $p < 0.05$.

Relative reduction in %Dis at age 3	Sample size
75%	46
50%	100
25%	390
Proportion with %Bx progression	
10%	36
20%	76
30%	208

Table 5: Sample sizes required for clinical trials designed with equally sized treatment and placebo arms, with 80% power at a significance level of $\alpha = 0.05$.

FIGURE CAPTIONS

Figure 1: A) Slice from inspiratory CT scan with overlaid grid. B) Slice annotated with PRAGMA-CF (inspiratory): red = bronchiectasis; yellow = mucous plugging; orange = bronchial wall thickening; pink = atelectasis; green = no abnormality seen. C) MinIP filtered slice from expiratory CT scan with overlaid grid. D) Slice annotated with PRAGMA-CF (expiratory): blue = trapped air, green = not trapped air.