ERJ Express. Published on June 18, 2009 as doi: 10.1183/09031936.00138508

Comparison of hyperpolarised ^{3}He MRI and HRCT in normal volunteers, patients with COPD and patients with alpha-1-antitrypsin deficiency – PHIL trial

Hyperpolarised 3-He MRI vs HRCT in COPD and normal volunteers – PHIL trial.

Edwin JR van Beek MD[1,7], Anja M Dahmen MD[2], Trine Stavngaard MD[3], Klaus K Gast MD[2], Claus Peter Heussel MD[2,5], Frank Krummenauer[8], Joerg Schmiedeskamp PhD[4,9], Jim M Wild PhD[1], Lise Vejby Søgaard PhD[3], Andreas E Morbach PhD[2], Laura M. Schreiber PhD[2], Hans-Ulrich Kauczor MD[2,6]

Academic Department of Radiology[1], University of Sheffield, Sheffield, United Kingdom

Department of Radiology[2], and Institute of Physics[4], Johannes Gutenberg University, Mainz, Germany

Danish Research Centre for Magnetic Resonance[3], Copenhagen University Hospital, Hvidovre, Denmark

Department of Radiology[5], Clinic for Thoracic Diseases at University of Heidelberg and Department of Radiology[6], German Cancer Research Center, Heidelberg, Germany

Department of Radiology[7], Carver College of Medicine, University of Iowa, Iowa City, USA

Clinical Epidemiology and Health Economy Unit[8], University Hospital Carl Gustav Carus, Dresden University of Technology, Dresden, Germany

Max Planck Institute for Polymer Research[9], Mainz, Germany

Address for correspondence:

Edwin JR van Beek MD PhD, Professor of Radiology, Carver College of Medicine, C-751 GH, 200 Hawkins Drive, Iowa City, Iowa 52242-1077, USA, e-mail: edwin-vanbeek@uiowa.edu

Abstract

Purpose

To apply hyperpolarized (HP) ^{3}He MRI to identify patients with chronic obstructive lung disease (COPD) from healthy volunteers and compare HP ^{3}He MRI findings with CT in a multicentre study. Quantitative measurements of HP ^{3}He MRI (apparent diffusion coefficient = ADC) and CT (mean lung density = MLD) were correlated with pulmonary function tests.

Material and Methods

A prospective three centre study enrolled 122 subjects with COPD (either acquired or genetic) and age matched never smokers. All diagnostic studies were completed in 94 subjects (52 with COPD; 13 with alpha-1-antitrypsin deficiency (A1AD); 29 healthy subjects; 63 men, 31 women; median age 62 years). The consensus assessment of radiologists blinded for other test results estimated non-ventilated lung volume (HP ^{3}He MRI) and percentage diseased lung (CT). Quantitative evaluation of all data for each center consisted of ADC (HP ^{3}He MRI) and MLD measurements (CT), and correlation with FEV_1/FVC indicating airway obstruction, and CO diffusion capacity (DLCO) indicating alveolar destruction.

Results

Using lung function tests as reference, regional analysis of HP ^{3}He MRI and CT correctly categorized normal volunteers in 100 and 97 %, COPD in 42% and 69% and A1AT in 69% and 85% of the cases, respectively. Direct comparison of HP ^{3}He MRI and CT revealed 23% of subjects with moderate/severe structural abnormalities had only mild ventilation defects. In comparison with lung function tests, ADC was more effective in separating COPD patients from healthy subjects than MLD ($p<0.001$ vs. 0.038). ADC measurements showed better correlation with DLCO than MLD ($r=0.59$ vs. 0.29).

Conclusion

Hyperpolarized ^{3}He MRI correctly categorizes patients with COPD and volunteers. It offers additional functional information, without the use of ionizing radiation whereas HRCT gives better morphological information. We showed the feasibility of a multcentre study using different MR systems.

Keywords: chronic obstructive lung disease, emphysema, hyperpolarized ^{3}Helium MRI, HRCT, ADC, MLD, DLCO, pulmonary function

Introduction

Chronic obstructive lung disease (COPD) is characterized by irreversible airflow obstruction [1]. The prevalence of COPD is increasing, largely as a result of the combined effects of smoking on an ageing population, and this realization has only become more urgent over the past decade [2,3]. Furthermore, the prevalence of this disease is affecting younger people, leading to early disability, high costs for healthcare and loss of economic contributions [3,4]. The impact of COPD on society as a whole is therefore highly significant [4], and it currently is the costliest disease within the western world, costing the UK National Health Service £2,576 million (€ 3.7 million) in 2000 [5].

Treatment options have been relatively limited. The foremost drive has been to reduce smoking, and prevention of the disease is an important step forward. However, the effects of such behavioural changes take more than 25 years to become visible; we are currently faced with the smoking history of patients, who are largely over the age of 45 years [3,5]. Broncho-dilatory and anti-inflammatory inhaled drugs are of limited value in lungs that are already severely damaged or destroyed. Lung volume reduction surgery has been shown to improve the quality of life of patients, but long-term survival effects have been disappointing and significant mortality and morbidity has been reported [6,7]. In the National Emphysema Treatment Trial, it was shown that selection of patients was possible using relatively simple HRCT based methods to demonstrate upper lobe predominance of disease [8]. Newer less-invasive methods of emphysema treatment via bronchoscopy are being tested [9,10].

In order for COPD patients to receive better treatment, both early identification and the linking of imaging with lung function are generally considered important. It is clear that these two diagnostic pathways offer different, complementary results. In clinical practice, most physicians rely on lung function tests, even though these are global measures and probably not sensitive enough to detect minor changes that may occur due to the disease process or in response to (novel) therapies. HRCT has been used to better define and quantify

both the distribution and extent of emphysema [8,11-15]. Indeed, quantification of emphysema has been shown to be quite accurate and reproducible, allowing use in the assessment of disease extent and possibly treatment response [13-15]. However, there are concerns over the use of ionizing radiation, particularly if scanning is to be performed at a younger age for early detection purposes or where longitudinal follow-up studies are required in chronic disease states, and therefore there is a potential application for MR-based techniques.

Proton MRI has (limited) capabilities in imaging chronic obstructive lung disease, but novel techniques are being tested that have additional contrast capabilities, and these have shown promising results [16-20]. Hyperpolarized ^{3}Helium MR imaging (HP ^{3}He MRI) has been developed over the past decade and has increasingly been introduced into clinical studies to assess the diagnostic performance in emphysema [21-25]. Although various techniques are under development, focus thus far has been on demonstration of qualitative or semi-quantitative ventilation distribution using a single breath-hold of a mixture of HP ^{3}He and Nitrogen gas and on assessment of "alveolar size" by measuring gas diffusivity (apparent diffusion coefficient). In particular, the apparent diffusion coefficient (ADC) measurements have shown to be sensitive for detection of micro-structural changes related to the most distal airways [26-29], including gravity effects [30], lung development [31] and ageing [32]. Most studies, however, have been relatively small and multi-centre studies have been rare [29]. Furthermore, the availability of HP ^{3}He gas has hampered wider evaluation of the technology and introduction of this technique into the clinical arena.

This multi-centre project aimed to assess the diagnostic capabilities of HP ^{3}He MRI in healthy volunteers and patients with acquired or genetic forms of COPD. A secondary aim was to compare the potentially useful quantitative parameters of HP ^{3}He MRI (apparent diffusion coefficient; ADC) with HRCT (mean lung density; MLD) and correlation of these parameters with results from pulmonary function tests (PFT; including FEV_1/FVC and CO

diffusion capacity (DLCO)). Finally, this study aimed to demonstrate the feasibility of performing multicentre studies with distant sites and a central 3-He gas production facility.

Materials and Methods

Study design

The study was performed in three centres, using a central gas production facility at the Institut für Physik, Mainz, Germany. Highly polarized ^{3}He gas was produced, shipped to the imaging site by road and/or air, and used for imaging at polarization levels ranging from 40% to 70% [33, 34]. This facility has obtained an EU manufacturing license and regulatory approval for distribution of ^{3}He gas for human imaging since the completion of this study. Individual approval was obtained from local Institutional Review Boards, and all patients gave informed consent prior to enrollment into the study.

All patient data were recorded at a central data processing centre. This allowed management of case-control subjects, according to age and gender distribution of target subjects with normal controls. All diagnostic tests were required to be performed within 1 week to reduce potential influence of COPD exacerbations or other respiratory illness.

Study population

This prospective case-control study aimed to recruit normal volunteers and subjects with COPD with an additional special subgroup of subjects with alpha-1-antitrypsin deficiency (A1AD). COPD was defined according to the guidelines of the European Respiratory Society as subjects with chronic chest symptoms with FEV1/FVC <70% predicted, FEV_1 <80% and reversibility of <12% and <200 ml, and asthmatics were specifically excluded from this study based on history or reversibility as above [34]. Subjects were restricted by age, only allowing participants who were older than 50 years. Smoking history did not serve as an inclusion criterion for patients. Normal volunteers were never

smokers (defined as fewer than 5 pack years and stopped at least 20 years prior to recruitment into the study), older than 50 years with FEV1>80% predicted and normal high resolution CT. Normal volunteers were recruited to match the age and gender of the patient group. Patients who underwent chest CT for other reasons and met the criteria for normal volunteers were also eligible for this group. Exclusion criteria were: pregnancy or breast feeding, myocardial infarction or stroke within the past 6 months, a history of pulmonary infection within prior 6 weeks and contraindication to MRI (pacemaker, ferromagnetic implants, etc.) as well as reversibility of airflow obstruction (as defined below).

Pulmonary function tests

Pulmonary function testing was performed using standard spirometry and body plethysmography according to the recommendations of the European Respiratory Society [35]. Patients underwent flow-volume reversibility testing 15 minutes after inhalation of 1.0 mg terbutaline. An increase in FEV1 by >12% and >200 ml was considered significant reversibility of airflow obstruction and resulted in exclusion from the study. An extensive series of measurement values was obtained, but we focused on the following main parameters: FEV1/FVC as an index of airway obstruction, and DLCO as an index of alveolar destruction. The latter was available in 83 of these 94 subjects.

COPD was categorized as follows: 1. No COPD as normal spirometry with or without chronic chest symptoms; 2. mild COPD as FEV1/FVC <70%, FEV1 >80% predicted with or without chronic chest symptoms (representing GOLD I); 3. moderate COPD as FEV1/FVC <70%, FEV1 between 30% and 80% predicted with or without chest symptoms (representing GOLD II and III); 4. severe COPD as FEV1/FVC <70%, FEV1 <30% predicted or presence of respiratory failure or clinical signs of right heart failure (representing GOLD IV) [1].

Magnetic Resonance Imaging

Ventilation distribution was evaluated using HP ^{3}He gas density distribution MRI using three 1.5 T MRI systems (2 Siemens Magnetom Vision, 1 Philips Eclipse) in the three study centres (Copenhagen, Mainz and Sheffield, respectively) tuned to the ^{3}He Larmor frequency of 48 MHz using specifically designed RF coils (2 Fraunhofer Gesellschaft, St. Ingbert, Germany, 1 Medical Advances, Milwaukee, USA). Patients were scanned in the supine position at breath-hold following the administration of 200-300 ml hyperpolarized ^{3}He gas through a face mask under control of a microprocessor controlled delivery device developed at one of the participating institutions [36]. The ^{3}He bolus was followed by additional volume of room air to circumvent anoxia and to reach full inspiration. The following sequences for morphologic 3-He imaging were used on the Philips Eclipse 1.5T MRI system (Sheffield): 18 s breath-hold sequence (FOV 42 cm, 128x128 matrix, TE/TR/flip angle 2.5 ms/ 7 ms/ 9 degrees, 19 coronal slices with a slice thickness of 10 mm). The Copenhagen and Mainz sites used Siemens Vision 1.5T systems: 12.5 s breath-hold (FOV 34x34 cm, 81x128 raw data matrix, TE/TR/flip angle 4.2 ms/11 ms/<10 degrees, 14 coronal slices with a slice thickness of 10 mm).

Diffusion imaging was performed to evaluate small airspace morphology using the following sequence (with some adaptation according to MR systems): a 12 s breath-hold sequence (FOV 47x47 cm, raw data matrix 64x128, TE/TR/flip angle 10.7 ms/16.1 ms/3 degrees, 3 transverse slices with a slice thickness of 20 mm at the level of the carina, 3 cm above and 5 cm below). The diffusion gradient has the following parameters: τ = 300 μs, $\delta = \Delta$ = 2300 μs, and G_{diff} = 12 mT/m, resulting in a *b*-value of b = 3.89 s/cm^2. The diffusion gradient was switched on along the x-, y- and z-axis to obtain ADC values in all three directions. From the images, a map of the mean apparent diffusion coefficient (ADC) in the axial plane was calculated as described previously [28]. A signal-to-noise ratio threshold of greater than 5 was accepted for inclusion into the study.

High resolution CT

For CT, a high-resolution algorithm was applied and 1mm axial slices at 10mm interval were reconstructed during full inspiration and full expiration. In normal volunteers, only 3 slices were acquired (3 cm above the carina, at the carina and 5 cm below the carina), while the entire chest was imaged in COPD and A1AD groups.

Image analysis

All data were de-identified and sent to a central image database at one of the centres, where entry criteria were verified. These images were presented to a panel of at least three chest radiologists from the three involved sites during four reading sessions. The radiologists were blinded to clinical details and evaluation was by consensus. The cases were read digitally in repetitive series of five randomly selected HP ^{3}He MRI investigations followed by five randomly selected HRCT studies at preset window setting (W/L 1000/-800).

For both HP ^{3}He MRI and HRCT the severity of ventilation impairment was visually estimated as a percentage of non-ventilated (HP ^{3}He MRI) or diseased lung (HRCT), respectively. For HP ^{3}He MRI, non-ventilated lung was estimated as the difference between the ventilated lung, which is clearly demonstrated by HP ^{3}He MRI, and the whole volume which is estimated from the presumed shape of the lungs within the thorax based on anatomic knowledge by the radiologists. For HRCT, diseased lung was defined as areas showing visually determined airway wall thickening and/or emphysematous destruction compared to the whole lung volume with estimation of affected lung segments. The HRCT disease extent was estimated based on the available slices. A real volumetric assessment was not possible due to restrictions of the imaging protocol. A semi-quantitative score was adopted for this process, ranging from normal to severe, which was applied for HP ^{3}He MRI and HRCT: 1. normal < 10%; 2. mild 11 - 30%; 3. moderate 31 - 60%; 4. severe >60%..

In addition quantitative evaluations were performed. For HP ^{3}He MRI, the mean ADC of the whole lung was calculated from the diffusion imaging, which was successfully obtained from 84 subjects [28]. At HRCT, mean lung density (MLD) was determined from 92 subjects, using previously described methods [37].

Statistical analysis

Visual analysis

The confirmatory analysis of the trial intended to assess the agreement between the parallel ^{3}He MRI and the CT findings based on respective binary classifications into "none or mild COPD" versus "moderate or severe COPD" findings. The descriptive agreement analysis therefore derived the relative frequencies of concordant and discordant parallel classifications; the significance evaluation of the latter was based on a McNemar test to evaluate the order of discordant findings. Furthermore Cohen's kappa coefficient (with asymptotic 95% confidence interval) was estimated to quantify the parallel classifications' agreement [38].

An extensive exploratory analysis then contrasted the underlying four stage visual ^{3}He MRI and CT evaluations in terms of relative frequencies; in addition the latter were contrasted with the patients' status as indicated by lung function assessments.

Quantitative analysis

A second stage of this exploratory analysis concentrated on the respective continuous parameterisation of lung function assessments, ^{3}He MRI and CT. Medians and quartiles for ADC and MLD were used for the description of these continuous endpoints; stratification for patients' health status and binary ratings was also based on medians and quartiles, graphical evaluations were based on non-parametric box whisker plots, accordingly.

Exploratory significance tests were summarized in terms of p-values, where $p<0.05$ was declared as an indication of locally statistical significance. The comparison of patient

samples along continuous endpoints was based on global Kruskal/Wallis tests and pair-wise two sample Wilcoxon tests. The comparison of samples along categorical endpoints was based on exact Fisher tests.

The pair-wise correlation between continuous endpoints was estimated by means of partial Spearman correlations controlling for patients' health status.

All numerical and graphical analyses were performed using the SPSS® software (release 12.0 for Windows®).

Sample size considerations

The overall investigation was intended to evaluate a fraction of 30% discordant rating patterns between the binary ^{3}He MRI and CT visual evaluations by means of a two-sided McNemar tests at the significance level 5%. To achieve a minimum power of 80% under the assumption of a rate difference of at least 15% in positive findings between MRI and CT, a minimum sample size of n=98 patients must be analysed.

Results

A total of 122 patients were enrolled into the study. Six patients were excluded as they did not comply with inclusion criteria on review. Three patients withdrew their consent during the course of the study. MRI was not completed in 15 participants. This was due to patient-related factors in 10 cases (claustrophobia 5, obesity 2, and unstable condition/dyspnea 3) and technical failure in 5 cases. In three MRI exams the obtained signal to noise ratio was considered insufficient to allow for evaluation (inadequate study). In one case CT was only available as a hardcopy, which neither allowed for sufficient image quality nor allowed for consistent image windowing. Finally, the data of 94 subjects were eligible for further evaluation with full analysis of MRI, CT and pulmonary function tests. The study population consisted of 52 patients with COPD (2 non-smokers = 4%), 13 with A1AD (12 non-smokers

= 92%), and 29 healthy subjects (all non-smokers). There were 63 men and 31 women, with a median age of 62 years (range 50-79 years). In 95% of subjects, the imaging procedures were performed on the same day, while PFT testing was done within one day in nearly all subjects.

No severe adverse reactions were observed and no medical interventions were required in any of the subjects studies. During the MRI investigation, pulse oxymetry measurements did not show any critical decrease.

Demographics of the study cohort, pulmonary function test results and their distribution across the three study centers are summarized in Table 1. The table also provides a comparison of the data between the recruited cohort of 116 subjects (small characters in brackets) and the sample actually eligible for further evaluation (n=94). As Table 1 illustrates, the necessary exclusion of subjects was not found to introduce a selection bias into the analysis cohort. Furthermore, the cofactors' distributions among the three study sites centers did not differ significantly or relevantly. However, as an effect of easier access to hyperpolarized ^{3}He gas, the radiology centre closest to the gas production centre recruited a slightly larger number of patients and healthy subjects than the other centres, as the logistics and organization of imaging was slightly more flexible.

COPD categorization according to GOLD criteria revealed that 28 of the 29 healthy non-smokers had normal lung function, whereas by our definition one subject had mild COPD. Of the 52 COPD patients 3 (6%) had mild, 39 (75%) moderate and 10 (19%) severe COPD, and of the 13 A1AD patients 1 (8%) had mild, 10 (77%) moderate, and 2 (15%) severe COPD.

Visual analysis

From HP ^{3}He MRI, 20 (69%) of the healthy non-smokers had normal ventilation distribution, and 9 (31%) had mild ventilation defects. Of the COPD and A1AD patients, this

distribution was 1 (2%) and 0 (0%) normal, 29 (55%) and 4 (31%) mild, 18 (35%) and 8 (61%) moderate and 4 (8%) and 1 (8%) severe ventilation impairment, respectively.

From HRCT, 20 (69%) of the healthy non-smokers had normal, 8 (28%) mild and 1 (3%) moderate pathologic findings. For the COPD group, 2 (4%), 14 (27%), 23 (44%) and 13 (25%) had normal, mild, moderate and severe HRCT abnormalities, while for A1AD patients these figures were 0 (0%), 2 (15%), 6 (46%) and 5 (39%), respectively.

For further evaluation, subjects were categorized in two groups as having normal or GOLD category 1 versus clearly impaired results (GOLD categories 2, 3 or 4) for PFT and normal or minimal vs. significant abnormalities on HP ^{3}He MRI and HRCT. Table 2 demonstrates the resulting fraction of subjects with clearly pathological findings in the respective diagnostic tests. Pulmonary function tests were used to define the pulmonary function status and classification of patients, and are therefore the reference method. The direct comparison of HP 3-He MRI with HRCT demonstrated that both were capable of identifying healthy subjects very well (Figure 1), but performed poorly in recognizing patients as having pathological results when pulmonary function tests were used as the reference. At the same time HRCT appeared slightly better than HP ^{3}He MRI in its ability to classify normal and diseased patients (Figures 1 – 2). To elucidate the role of both modalities a direct comparison was performed. Significantly more (McNemar p=0.002) HP ^{3}He MRI (67%) studies were categorized as “none or mild” disease based on the amount of ventilatory abnormalities with corresponding CT findings categorized as “none or mild” in 44% and “moderate or severe” in 23%; conversely both tests demonstrated “moderate or severe” disease in 28% and CT was relatively normal in 5% of patients with “moderate or severe” He-3 ventilation defects. Accordingly the overall agreement analysis between CT and MRI findings resulted in a kappa estimate of 0.43 (95% confidence interval 0.38 – 0.48) indicating moderate agreement between the binary ratings. Further examples of direct comparison between ^{3}He MRI and CT are demonstrated in Figures 3-6.

Table 3 demonstrates the correlation between CT and MRI as an overall assessment in subjects with COPD (Table 3A) and in normal subjects (Table 3B). Clearly, there is a shift towards normal as expected, but also for both categories the actual agreement between the two imaging modalities is 18/52 (35%) in the COPD group and 19/29 (66%) in the healthy group. Upon closer inspection, 8/52 (15%) and 1/29 (3%) of subjects have a disagreement by greater than one category, respectively.

Quantitative analysis

The second stage of the evaluation was based on quantitative readouts of both imaging modalities as Figure 7 illustrates the ADC measurement from HP ^{3}He MRI, and MLD from HRCT. ADC measurements provided a statistically significant differentiation (Wilcoxon $p<0.001$) between healthy subjects (median 0.17 cm^2/s) versus COPD patients (median 0.28 cm^2/s) and A1AD patients (median 0.29 cm^2/s), respectively. There was no significant difference between COPD and A1AD patients. The MLD distributions, however, hardly contrasted healthy subjects (median MLD -855 HU) from COPD patients (median MLD -865 HU; Wilcoxon $p=0.038$) from a clinical perspective (Figure 5b); on the other hand A1AD patients showed clearly lower values (median MLD -892 HU; Wilcoxon $p<0.001$ versus healthy and COPD subjects, respectively). After adjustment for pulmonary function test status, partial Spearman correlation coefficients demonstrated moderate correlations between FEV1/FVC as an indicator of airway obstruction with both ADC and MLD ($r = 0.50$ and 0.52, respectively), and between DLCO as an indicator of alveolar destruction with ADC and MLD ($r = 0.59$ vs. 0.29, respectively).

Discussion

The results of this multicentre study, the largest of its kind to date, showed the feasibility of performing this rather complex study in a multinational, multicentre fashion using a central production facility. This obviously is an important step forward towards the application of ^{3}He MRI.

The study demonstrated that there is a significant difference in information obtained from PFT and HP ^{3}He MRI compared to HRCT. In particular, it appears that ADC measurements using HP ^{3}He MRI are more closely correlated to functional information (DLCO) obtained through routine pulmonary function tests and CO diffusion capacity assessments. In addition, there are significant differences between ventilation distribution images and HRCT images. These observations could be explained by the fundamental differences between HRCT and HP ^{3}He MRI imaging. The latter relies on gas inflow into the lungs for visualization of the airways and the size of the (distal) pulmonary airspaces [39] and therefore is closer in nature to the breathing manoeuvres as used in pulmonary function tests. As the visual assessment of both HP ^{3}He MRI and CT was performed using categorical scores, and not continuous variables, comparison is by definition a little more difficult. However, quantitative assessment using ADC and MLD are continuous variables, and therefore can be evaluated in greater detail.

Although not affecting the direct comparison between HRCT and MRI, one needs to also consider that these imaging modalities allow for regional information. However, as a comparison with pulmonary function tests were necessary (as the reference method), this regional information was lost as an averaging process was needed to allow for statistical comparison. Healthy lung parenchyma compensates (at least in part) for lack of parenchymal

function in diseased lung states, usually by compensatory hyperinflation and/or redirection of perfusion to enhance ventilation-perfusion matching and gas exchange mechanism. Regional functional HP ^{3}He MRI is able to distinguish some of the compensatory mechanisms, demonstrating redistribution of ventilation in the presence of ventilation defects. Thus, the dichotomy that is forced onto these tests by comparison with pulmonary function tests and biostatistical sample size calculation, as well as the arbitrary classification of level of disease based on the visual estimates, are likely resulting in under and over reporting of abnormalities in the subgroups that were studied using imaging methods. Thus, when combining the differences between the imaging and pulmonary function tests, one may argue that pulmonary function tests are (by their very nature) a sub-standard reference method in the assessment of novel imaging techniques of the lung.

This study aimed to compare two imaging modalities in a cohort of patients, who were grouped and defined according to pulmonary function tests. The problem arises that the imaging tests may be more sensitive to (regional) changes in the lung than the global results obtained by PFT; and this discrepancy obviously means that a comparison will be skewed from the outset. Nevertheless, there is significant correlation for both methods, thus allowing for both tests to be of use in the assessment and distribution analysis of emphysema. This information may become relevant for the planning and guidance of regional emphysema treatment, such as interventional bronchoscopic valve placement and other new management strategies [10]. Furthermore, new pharmacological treatment trials may benefit from regional functional tools, as these appear to be more sensitive and may therefore reduce the sample size as outcome may be based on these surrogate measures.

There is some validity to separate normal/mild COPD from moderate/severe COPD as an outcome parameter. First, the GOLD criteria have been shown to be reproducible in multiple centres, and they have been widely adapted. Second, treatment will very likely be different, with medical therapy predominantly of use in earlier disease, while surgical or other

interventional treatment will be sought in more severe emphysema. Therefore, any imaging test should be capable of identifying those that require medical treatment versus those that need surgical or interventional treatment. However, one needs to realize that the chosen analysis with global values was required in order to compare HP ^{3}He MRI and HRCT with the accepted reference standard (pulmonary function tests). This limitation may not be as critical once these tests are applied for clinical practice or for assessment of treatment response, where they can serve as independent (and regional) outcome measures. In addition, the trade-off between sensitivity versus specificity is largely driven by the potential for treatment in those patients that are diagnosed early. Thus, as long as there is no effective early treatment available, this is a decision that cannot currently be made.

The correlation between HRCT and ventilation images was probably influenced by the way in which the visual assessment was designed. It has become clear that visual assessment can be less accurate than computer-based assessment, both for MRI [40] and for CT [13-15]. The study protocol did not intend for such an analysis to take place. In retrospect, the additional performance of a proton MRI sequence prior to HP ^{3}He MR imaging would have enabled the performance of a later proposed method [40]. This method uses a proton MRI sequence as a mask for the subsequent subtraction of signal from HP ^{3}He MRI, and would have been able to demonstrate signal-based ventilated lung volumes bases on absolute pixel values [40]. In addition, with the advancement of CT scanner technology and the subsequent changes in CT protocols, it is now possible to allow full volumetric lung assessment with histogram analysis and determination of extent of emphysema based on various density cut-off values [15]. A recent study in an emphysema-induced canine model compared volumetric CT with 3 mm slice thickness and ^{3}He diffusion MRI and demonstrated very good correlation between the two modalities [41]. A further study in asthmatic patients also demonstrated good correlation between MDCT findings of airway morphology and expiratory air trapping and the presence of ventilation defects [42]. This discrepancy of our relative lack of agreement

compared to this literature may well be related to the fact that we used traditional HRCT methods (with axial 1 mm slices and 1 cm gaps), thus reducing the ability to obtain better comparative images.

A second issue is that we were forced to compare axial CT slices with coronal MRI slices, again due to the traditional HRCT methodology of our study. It was feasible to obtain axial reconstructions of MRI (as these were volumetrically acquired images), but this comparison was not perfect. Nevertheless, by trying to score pathological findings based on segmental involvement, a reasonable correlation would have been expected if the two tests were demonstrating the same pathology.

The apparent diffusion coefficient appears to be most closely related to functional parameters, in particular DLCO. This is not unexpected, as it is a measurement of microstructure of the lung, which is totally dependent on ventilation taking place to the regions under review. As a result, mean ADC will probably underestimate the amount of destroyed lung tissue, as inflow obstruction in COPD will prevent ADC measurement, and some of the most destroyed emphysematous lung areas will therefore be excluded from the analysis. In spite of this limitation, the ADC was more closely related to pulmonary function tests than HRCT. Ongoing analysis is taking place to assess regional differences between the two imaging methods. One should, however, note that the ADC depends on the level of inspiration which was not spirometrically controlled in the present study. A further point of note is that the measured ^{3}He ADC values are dependent upon pulse sequence timing of the gradients and their associated diffusion weighted b-values [43]. In this study relatively strong b-values $b = 3.89$ s/cm^2 were used to achieve a strong diffusion weighting and, thus, a high sensitivity for regional differences in the ADC values. This approach, however, results in a complete loss of MRI signal in the trachea and grossly enlarged alveolar spaces. In the ADC pulse sequence, diffusion gradients were applied subsequently on all axes. An earlier study which evaluated a subset of the volunteer data presented here has shown that in normals no

differences are observed between the ADC values obtained in the three directions [29]. Moreover, it has been shown that no differences between the ADC values are observed if an identical pulse sequence is used on the same MRI scanner model at two different study centres

In the majority of other studies published in COPD with ^{3}He MRI [27,32]] a lower b-value of b = 1.6 s/cm^2 on a single axis was used. Care should be taken when making a direct comparison of our ADC results with these smaller studies as the higher b-values used here may give a lower ADC [44]. In future studies some international consensus on standardisation of b-value for ADC measurement would be worthwhile. Nevertheless for inter patient comparison in this multi-centre study the b-values used and ADC values obtained are consistent between sites.

Although this study underwent an a priori statistical sample size calculation to achieve sufficient power of the overall evaluation, it must be emphasized, that some of the above results are due to exploratory analyses and therefore may suffer from loss in statistical power. During the planning phase of the trial, the investigation was not designed to thoroughly stratify results for the patients' disease status (smoker, COPD, alpha), but should rather derive an "overall" agreement statement between CT and ^{3}He-MRI based diagnostic ratings. As a consequence, the overall trial's sample size was designed for an overall agreement analysis at an 80% statistical power level without any multiplicity consideration. However, during initial descriptive analyses, severe sub sample heterogeneity according to the patients' disease status was observed. A stratification of all clinically relevant results was therefore found necessary and provided throughout the evaluation. In fact, the resulting sub sample evaluations will suffer from a loss in statistical power, which cannot be corrected for by formal multiplicity correction, but could have only been taken into account for by interim sample sizes correction. As these findings were due to descriptive, in some sense exploratory evaluations, the authors decided to display the numerical data in a corresponding manner. For example,

Table 2 therefore only provides stratified point estimates instead of formally multiplicity adjusted confidence intervals, which again would have afforded larger sample sizes in the first place. Should a study be aimed at a four-class comparison of these diagnostic tests, a study population of over 200 subjects would be necessary and this was too much to achieve in the setting of this multicenter study using this new technology.

Finally, although not formally assessed in the present study, a possible issue that needs further evaluation is the repeatability of ^{3}He MRI, not only in normal subjects, but also in those with asthma and COPD. At least two recent studies have directly assessed this issue. One study in asthmatic patients demonstrated that ventilation defects in asthma patients tending to be quite persistent in spite of optimal treatment, even though a decrease in the extent of the ventilation defects was observed in most patients [45]. A further study in patients with COPD demonstrated excellent same-day and 1 week follow-up repeated measurements for ADC measurements, whereas measurements aimed at ventilation distribution were less reproducible [46].

In conclusion, the current study findings suggest that HRCT is capable of demonstrating the morphological (airways) characteristics of COPD (such as bronchial wall thickening and bronchiectasis). However, HP ^{3}He MRI appears slightly more closely related to functional measures, and complements HRCT with its capability to quantify ventilation within the lung regions, as well as apply diffusion measurements to allow for micro-structural changes on a spatial scale smaller than the resolution of HRCT. The techniques thus offer quite different, complementary information, which each may be useful for treatment planning in different contexts. The exact roles of these two imaging techniques need to be established using management studies.

Acknowledgement

The following people participated in the conduct of this study:

Johannes Gutenberg University, Institut für Physik (Ernst Otten, Werner Heil and Frank Filbir) and Department of Anesthesiology (Frank Knitz and Norbert Weiler), Mainz, Germany;

Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital (Lars Hanson, Karin Markenroth, Jann Mortensen, Asger Dirksen, Anne Kiil Berthelsen, Ejvind Frausing and Joergen Vestbo), Hvidovre, Denmark;

University of Sheffield, Departments of Anesthesiology (Gary Mills), Respiratory Medicine (Rod Lawson) and Academic Radiology (Sally Fleming, Stan Fichele, Neil Woodhouse, Andrew Swift, Martyn Paley and Paul Griffiths), Sheffield, UK.

Funding for this study was provided by the European Union, Framework 5, Polarized Helium Imaging of the Lung project (IST-2000-31559).

References

1. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. (2001) Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO. Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med 163:1256-1276.
2. Murray CJ and Lopez AD. (1997) Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. Lancet 349: 1498-504.
3. Lopez Ad, Shibuya K, Rao C, et al. (2006) Chronic obstructive pulmonary disease: current burden and future projections. Eur Respir J 27:397-412.
4. Chapman KR, Mannino DM, Soriano JB, et al. (2006) Epidemiology and costs of chronic obstructive pulmonary disease. Eur Respir J 27:188-207.
5. British Thoracic Society. The burden of lung disease. London, BTS, 2001.
6. Fishman A, Martinez F, Naunheim K, et al. (2003) A randomized trial comparing LVRS with medical therapy for severe emphysema. N Engl J Med 348:1059-2073.
7. Cooper JD, Patterson GA, Sundaresan RS, et al. (1996) Results of 150 consecutive bilateral lung volume reduction procedures in patients with severe emphysema. J Thorac Cardiovasc Surg 112:1319-1330.
8. Naunheim KS, Wood DE, Krasna MJ, et al. (2006) Predictors of operative mortality and cardiopulmonary morbidity in the National Emphysema Treatment Trial. J Thorac Cardiovasc Surg 131:43-53.
9. Wood DE, McKenna RJ Jr, Yusen RD, et al. (2007) A multicenter trial of an intrabronchial valve for treatment of severe emphysema. J Thorac Cardiovasc Surg 133:65-73.

10. Strange C, Herth FJ, Kovitz KL, et al. (2007) Design of the endobronchial valve for emphysema palliation trial (VENT): a non-surgical method of lung volume reduction. BMC Pulm Med 7:10.

11. Sanders C, Nath PH, Biley WC. (1988) Detection of emphysema with computed tomography. Correlation with pulmonary function test and chest radiography. Invest Radiol 23: 262.

12. Klein JS, Gamsu G, Webb WR, Golden JA, Muller ML. (1992) High-Resolution CT diagnosis of emphysema in symptomatic patients with normal chest radiographs and isolated low diffusing capacity. Radiology 182: 817-21

13. Bankier AA, de Maertelaer V, Keyzer C, Gevenois PA. (1999) Pulmonary emphysema: subjective visual grading versus objective quantification with macroscopic morphometry and thin-section CT densitometry. Radiology 211:851-858.

14. Uppaluri R, Mitsa T, Sonka M, Hoffman EA, McLennan G. (1997) Quantification of pulmonary emphysema from lung computed tomography images. Am J Respir Crit Care Med 156:248-254.

15. Hoffman EA, Simon BA, McLennan G. (2006) State of the Art. A structural and functional assessment of the lung via multidetector computed tomography: phenotyping chronic obstructive pulmonary disease. Proc Am Thorac Soc 3:519-532.

16. Quintana HK, Cannet C, Zurbruegg S, et al. (2006) Proton MRI as a noninvasive tool to assess elastase-induced lung damage in spontaneous breathing rats. Magn Reson Med 56:1242-1250.

17. Morino S, Toba T, Araki M, et al. (2006) Noninvasive assessment of pulmonary emphysema using dynamic contrast-enhanced magnetic resonance imaging. Exp Lung Res 32:55-67.

18. Ley-Zaporozhan J, Ley S, Kauczor HU. (2007) Proton MRI in COPD. COPD 4:55-65

19. Iwasawa T, Takahishi H, Ogura T, et al. (2007) Correlation of lung parenchymal MR signal intensity with pulmonary function tests and quantitative computed tomography (CT) evaluation: a pilot study. J Magn Reson Imaging Oct 29; [Epub ahead of print].

20. Swift AJ, Woodhouse N, Fichele S, et al. (2007) Rapid lung volumetry using ultrafast dynamic magnetic resonance imaging during forced vital capacity maneuver: correlation with spirometry. Invest Radiol 42:37-41.

21. Kauczor HU, Hofmann D, Kreitner KF, et al. (1996) Normal and abnormal pulmonary ventilation: visualization at hyperpolarized He-3 MR imaging. Radiology 201:564-568.

22. MacFall JR, Charles HC, Black RD, et al. (1996) Human lung air spaces: potential for MR imaging with hyperpolarized He-3. Radiology 200:553-558.

23. De Lange EE, Mugler JP, Brookeman JR, et al. (1999) Lung air spaces: MR imaging evaluation with hyperpolarized 3He gas. Radiology 210:851-857.

24. Van Beek EJR, Wild JM, Schreiber W, Kauczor HU, Mugler III J, de Lange EE. (2004) Functional MRI of the lung using hyperpolarized 3-Helium gas. J Magn Reson Imaging 20:540-554

25. Guenther D, Eberle B, Hast J, et al. (2000) (3)He MRI in healthy volunteers: preliminary correlation with smoking history and lung volumes. NMR Biomed 13:182-189.

26. Saam B, Yablonskiy DA, Kodibagkar VD et al. (2000) MR imaging of diffusion of ^{3}He gas in healthy and diseased lungs. Magnetic Resonance Medicine 44: 174-179.

27. Salerno M., De Lange EE, Altes TA et al. (2002) Emphysema: hyperpolarized helium 3 diffusion MR imaging of the lungs compared with spirometric indexes--initial experience. Radiology 222: 252-60.

28. Morbach AE, Gast KK, Schmiedeskamp J, et al. (2005) Diffusion-weighted MRI of the lung with hyperpolarized helium-3: a study of reproducibility. J Magn Reson Imaging 21:765-774.

29. Schreiber WG, Morbach AE, Stavngaard T, et al. (2005) Assessment of lung microstructure with magnetic resonance imaging of hyperpolarized Helium-3. Respir Physiol Neurobiol 25:23-42.

30. Swift AJ, Wild JM, Fichele S, et al. (2005) Emphysematous changes and normal variation in smokers and COPD patients using diffusion 3He MRI. Eur J Radiol 54:352-358.

31. Altes TA, Mata J, de Lange EE, Brookeman JR, Mugler JP 3rd. (2006) Assessment of lung development using hyperpolarized helium-3 diffusion MR imaging. J Magn Reson Imaging 24:1277-1283.

32. Fain SB, Altes TA, Panth SR, et al. (2005) Detection of age-dependent changes in healthy adult lungs with diffusion-weighted 3He MRI. Acad Radiol 12:1385-1393.

33. Wild JM, Schmiedeskamp J, Paley MNJ, et al. MR imaging of the lungs with hyperpolarized 3-Helium transported by air. Phys Med Biol 2002;47:N185-190.

34. Van Beek EJR, Wild JM, Schmiedeskamp J, et al. (2003) Hyperpolarized 3-Helium MR imaging of the lungs: testing the concept of a central production facility. Eur Radiol 13:2583-2586.

35. Celli BR, MacNee W. ATS/ERS Task Force. (2004) Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. Eur Respir J 23:932-946.

36. Eberle B, Weiler N, Markstaller K, et al. (1999) Analysis of intrapulmonary O2 concentration by MR imaging of inhaled hyperpolarized helium-3. J Appl Physiol 87:2043-2052.

37. Kauczor HU, Hast J, Heussel CP, Schlegel J, Mildenberger P, Thelen M. (2002) CT attenuation of paired HRCT scans obtained at full inspiratory/expiratory position: comparison with pulmonary function tests. Eur Radiol 12:2757-2763.

38. Krummenauer F. (2006) The comparison of clinical imaging devices with respect to parallel readings in both devices. Eur J Med Res 11:119–22.

39. Stavngaard T, Søgaard LV, Mortensen J, et al. Hyperpolarized 3He MRI and 81mKr SPECT in chronic obstructive pulmonary disease. Eur J Nucl Med Mol Imaging 2005;32:448-457.

40. Woodhouse N, Wild JM, Paley MNJ, et al. (2005) Combined Helium-3/proton MRI measurement of ventilated lung volumes in smokers compared to never-smokers. J Magn Reson Imaging 21:365-369

41. Tanoli TS, Woods JC, Conradi MS, et al. (2007) In vivo lung morphometry with hyperpolarized 3He diffusion MRI with induced emphysema: disease progression and comparison with computed tomography. J Appl Physiol 102:477-484.

42. Fain SB, Gonzalez-Fernandez G, Peterson ET, et al. (2008) Evaluation of structure-function relationships in asthma using multidetector CT and hyperpolarized He-3 MRI. Acad Radiol 15:753-762.

43. .Fichele S, Paley MN, Woodhouse N, Griffiths PD, van Beek EJ, Wild JM. Finite-difference simulations of 3He diffusion in 3D alveolar ducts: comparison with the "cylinder model". Magn Reson Med 2004;52:917-920.

44. Yablonskiy DA, Sukstanskii AL, Leawoods JC, et al. (2002) Quantitative in vivo assessment of lung microstructure at the alveolar level with hyperpolarized 3He diffusion MRI. Proc Nat Acad Sci (NY) 99:3111-3116.

45. De Lange EE, Altes TA, Patrie JT, et al. (2009) Changes in regional airflow obstruction over time in the lungs of patients with asthma : evaluation with 3He MR imaging. Radiology 250:567-575.

46. Mathew L, Evans A, Ouriadov A, et al. (2008) Hyperpolarized 3He magnetic resonance imaging of chronic obstructive pulmonary disease: reproducibility at 3.0 Tesla. Acad Radiol 15:1298-1311.

Legends

Table 1

Socio-demographic characteristics and lung function categorization of 116 subjects (small characters in brackets) and the sample actually eligible for further evaluation (n=94) across study sites. Medians for continuous endpoints and relative frequencies for categorical endpoints are given. The p-values derived from overall exact Fisher tests and Kruskal/Wallis tests across study sites confirm that the necessary exclusion of subjects did not introduce a selection bias into the analysis cohort. Furthermore, the cofactors' distributions among the three study sites centres were similar.

Table 1

a) Sociodemographics of the recruited cohort of 116 subjects, who all met the inclusion criteria

	n	median age [years]	Women	samples			smokers	VC [l]	FEV1 [l]	FEV1/VC [%]	DLCO [mmol/min/kPa]
				COPD	A1AD	Healthy					
Centre 1	57	62	16	23	12	22	23	3.53	1.69	48	5.89
Centre 2	24	66	11	16	2	6	14	3.32	1.49	40	5,05
Centre 3	35	62	11	23	3	9	24	2.88	1.67	58	5,11

b) Sociodemographics of the cohort of 94 subjects, who were eligible for further evaluation

	n	median age [years]	Women	samples			smokers	VC [l]	FEV1 [l]	FEV1/VC [%]	DLCO [mmol/min/kPa]
				COPD	A1AD	Healthy					
Centre 1	43	62	12	18	10	15	18	3.53	1.67	46	5.92
Centre 2	19	66	8	13	1	5	12	3.32	1.58	39	4,60
Centre 3	32	62	11	21	2	9	21	2.88	1.72	58	4,84
P (sites)		0.430	0.537	0.086	0.085	0.115	0.442	0.040 *	0.359	0.430	0.537

* Centre 3 appears to have included slightly more severely affected COPD cases compared to the other centres.

Table 2

Categorization of the subjects with clear pathological findings in the three diagnostic tests' groups as percentage of the COPD, A1AD and healthy cohorts.

Table 2

	% identified as mild, moderate or severe (no. in brackets)				% identified as moderate or severe (no. in brackets)		
	Pulmonary function test	HP 3-He MRI	CT		Pulmonary function test	HP 3-He MRI	CT
Healthy N=29 (FPR)	3.4 (1)	31 (9)	31 (9)		0 (0)	0 (0)	0 (0)
COPD N=52 (DR)	100 (52)	98 (51)	96 (52)		94 (49)	42 (22)	69 (36)
A1Ad N=13 (DR)	100 (13)	100 (13)	100 (13)		69 (9)	69 (9)	85 (11)

FPR=False-positive rate (ie 1-specificity)
DR= detection rate (ie sensitivity)

Table 3

Cross correlation of CT and HP ^{3}He MRI for the 52 subjects with COPD (a) and the 29 healthy subjects (b), demonstrating imperfect agreement between the two imaging modalities.

Table 3: Cross correlation of CT and HP 3-He MRI for the 52 subjects with COPD (a) and the 29 healthy subjects (b), demonstrating imperfect agreement between the two imaging modalities.

A.
52 patients with COPD

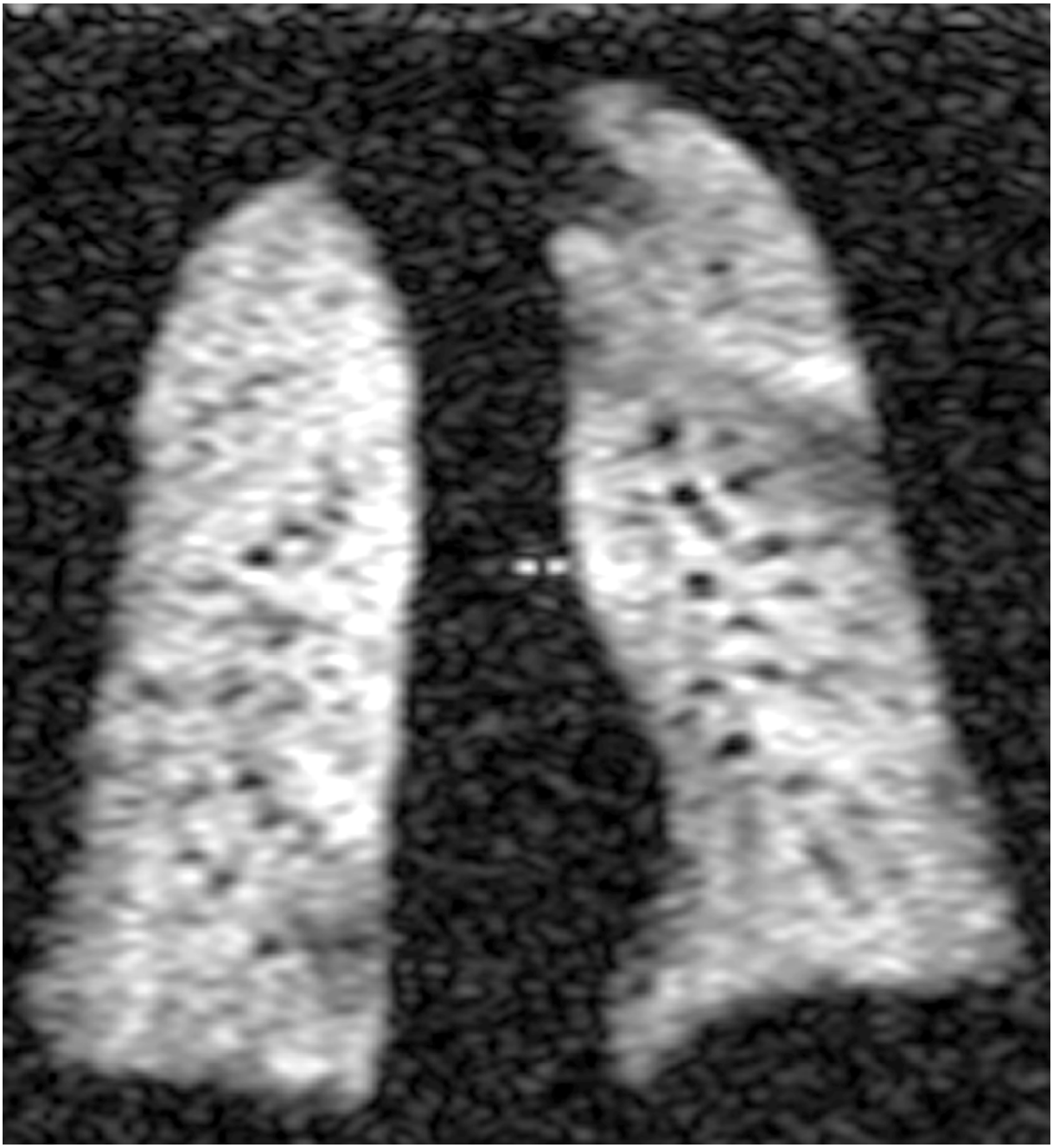
	HP 3-He MRI			
CT	Normal	Mild	Moderate	Severe
Normal	0	2	0	0
Mild	0	9	4	1
Moderate	1	12	8	2
Severe	0	6	6	1

B.
29 healthy subjects

	HP 3-He MRI			
CT	Normal	Mild	Moderate	Severe
Normal	15	5	0	0
Mild	4	4	0	0
Moderate	1	0	0	0
Severe	0	0	0	0

Figure 1:

62 year-old male healthy subject with normal lung function (FEV 114%); HP ^{3}He MRI (a) normal (9% non-ventilated lung) and CT (b) normal (8% diseased lung)



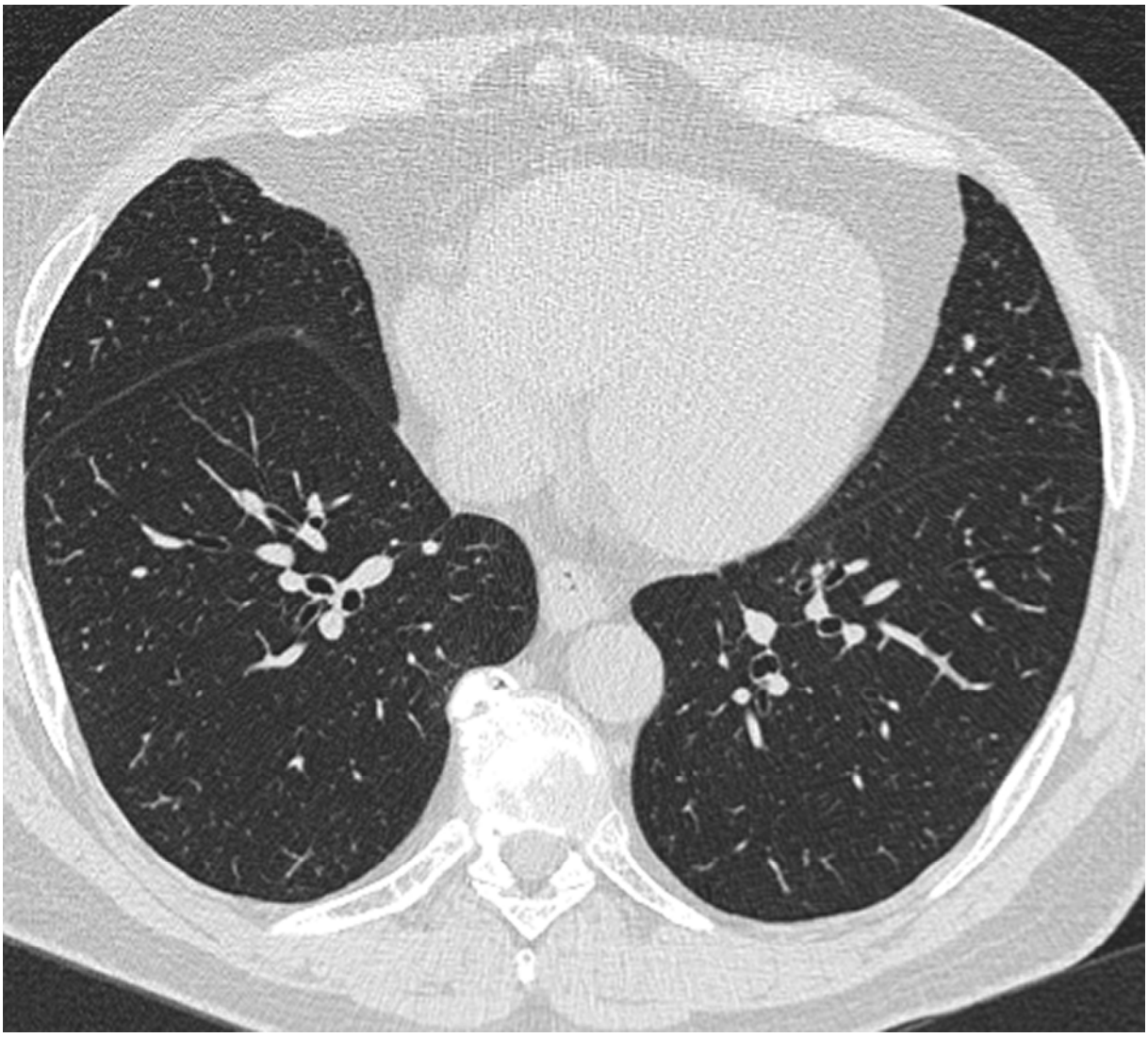
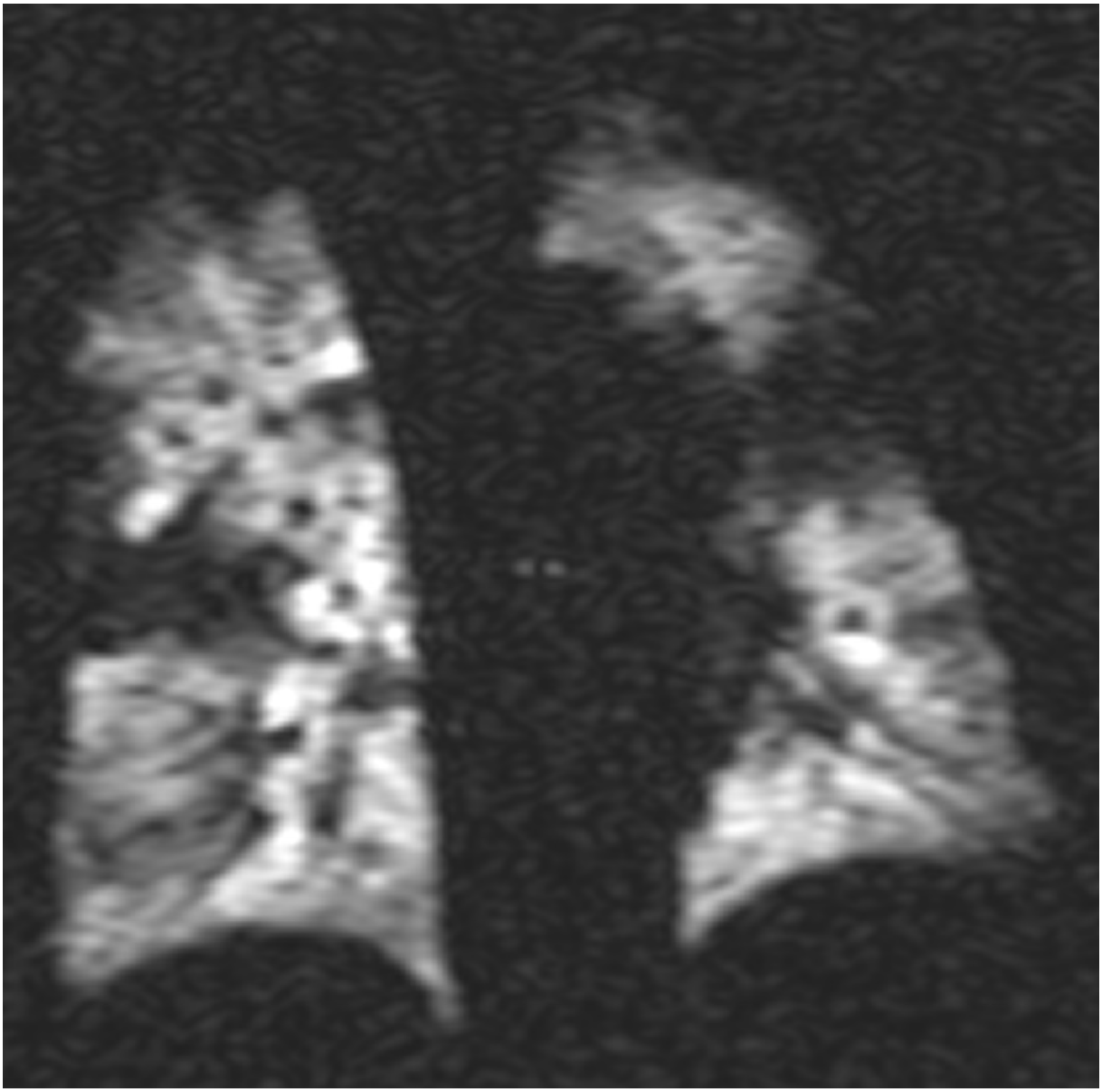


Figure 2:

69 year-old male healthy subject with normal lung function (FEV 88%); HP ^{3}He MRI (a) grade mild (20% non-ventilated lung) and CT (b) grade mild (15% diseased lung)



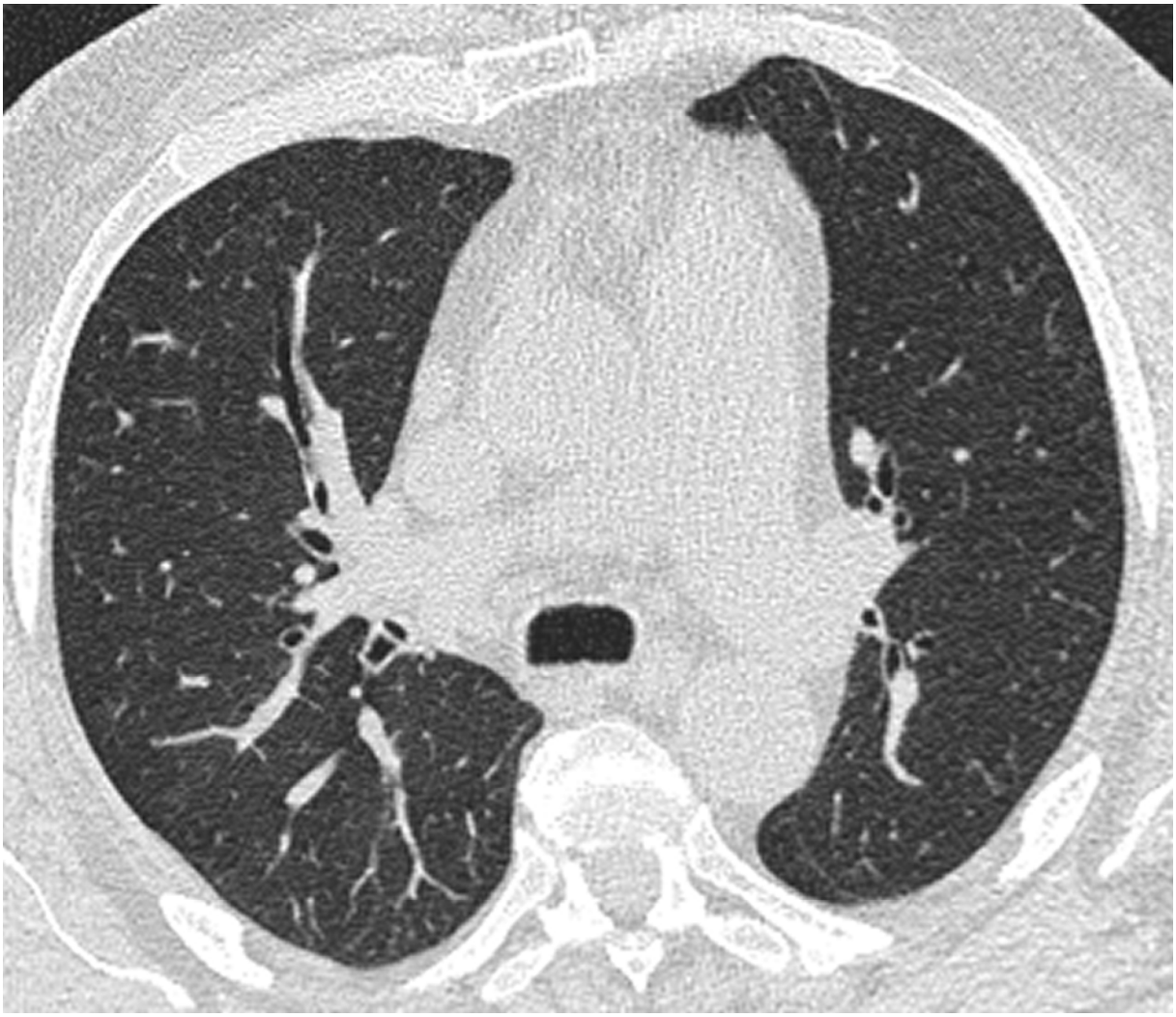
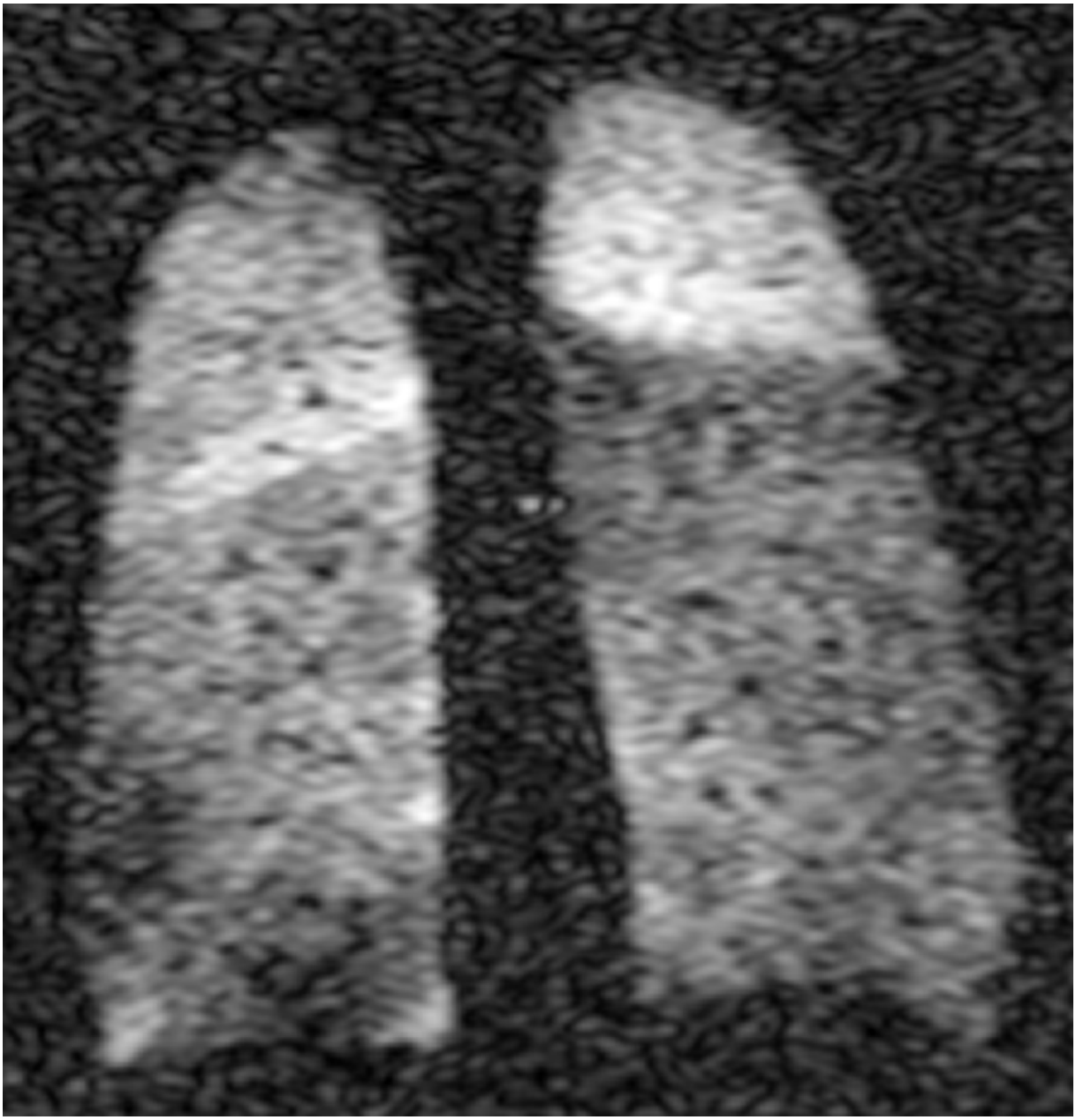


Figure 3:

60 year-old male patient with COPD grade moderate (FEV 75%); HP ^{3}He MRI (a) grade moderate (40% non-ventilated lung) and CT (b) normal (2% diseased lung). This case demonstrates that ventilation abnormalities don't necessarily coincide with morphology of the lung.



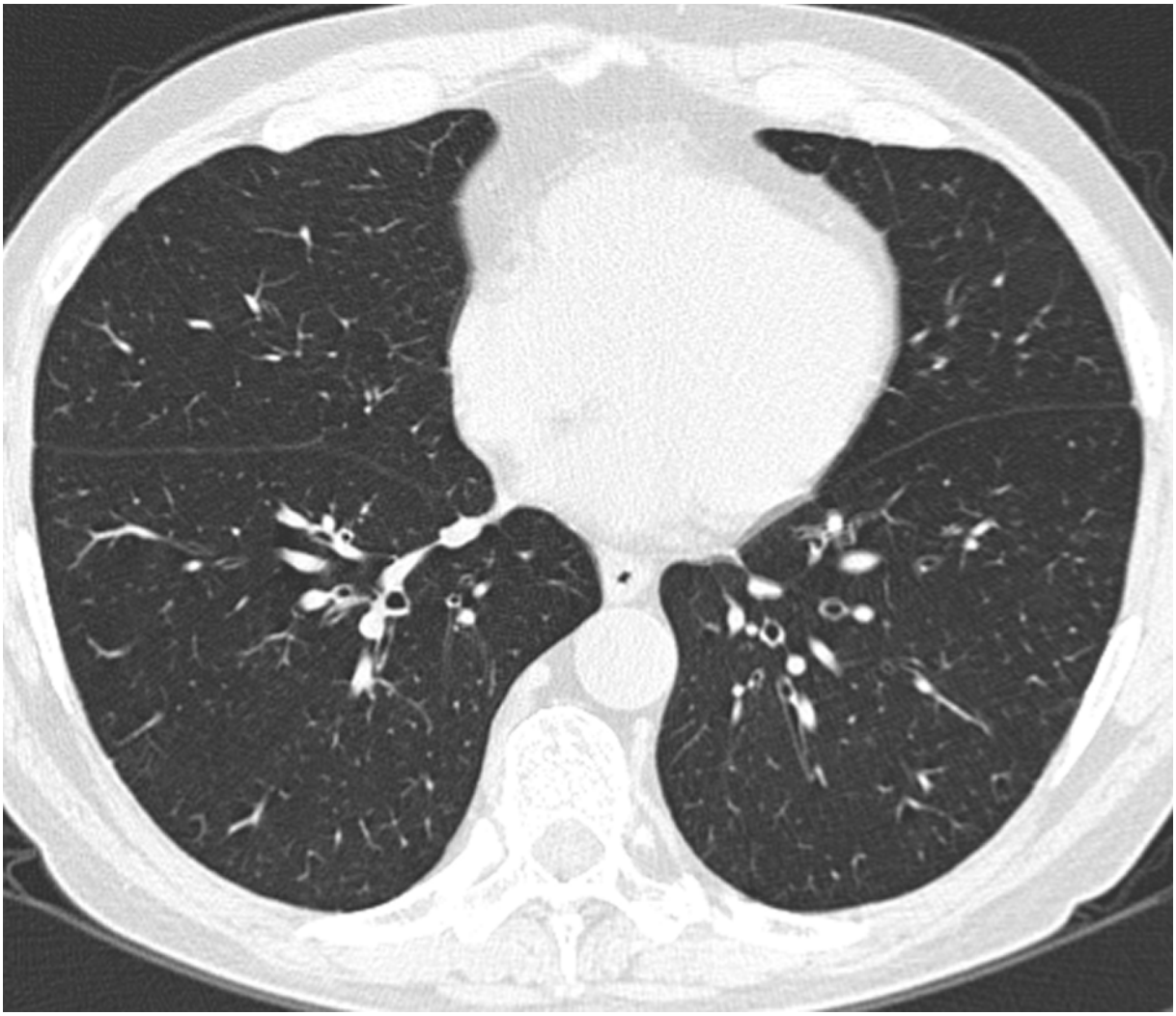
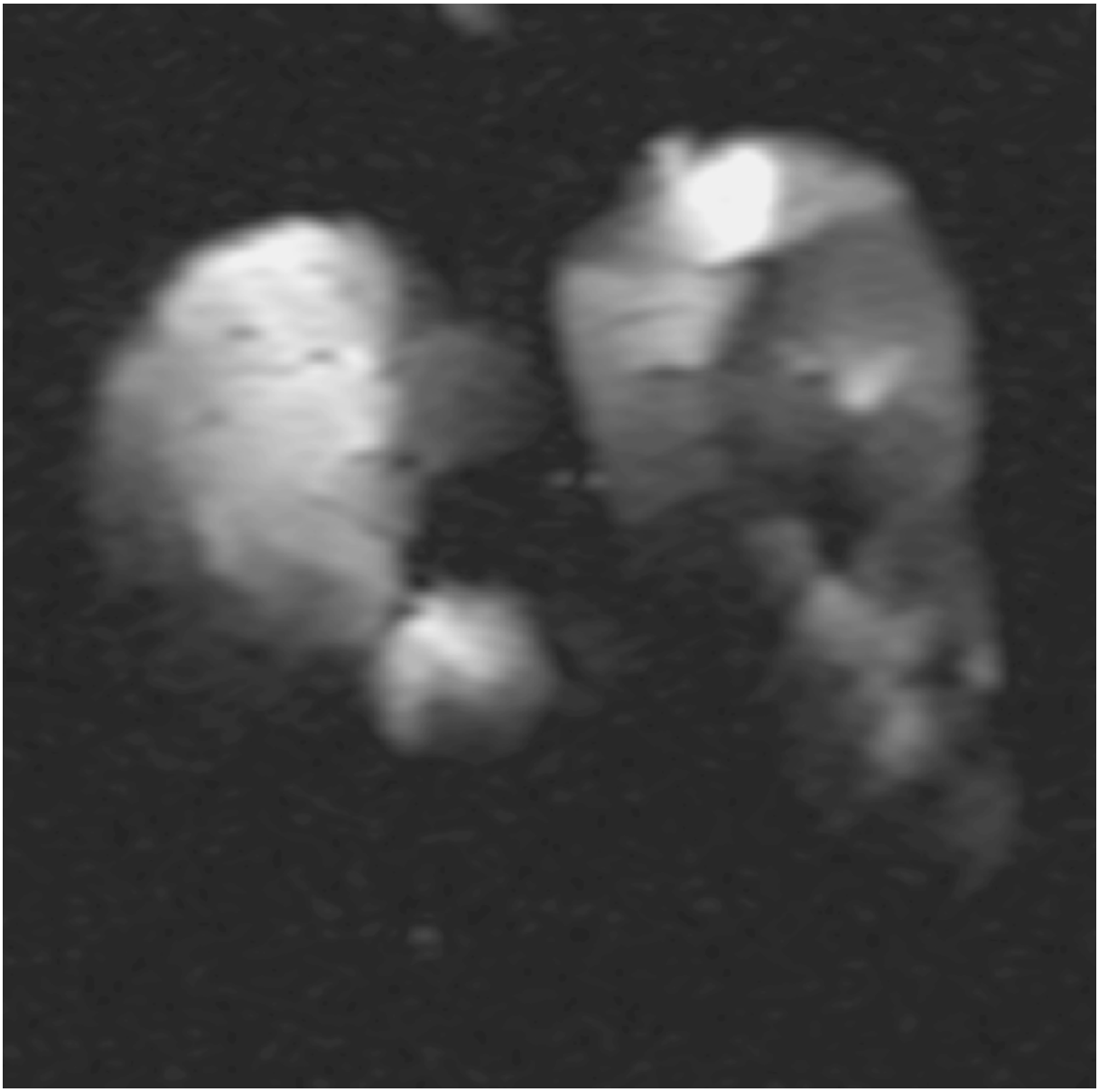
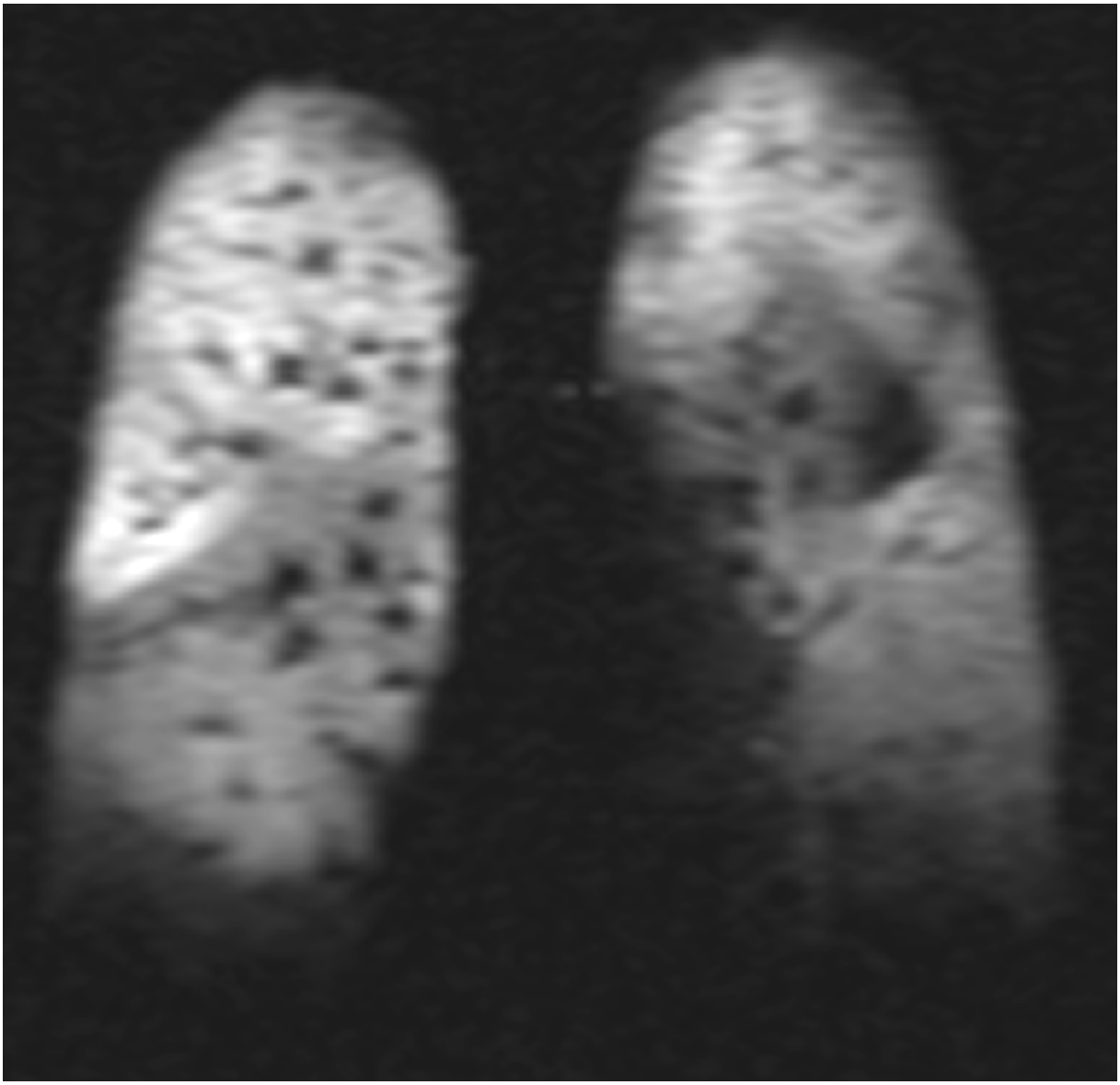


Figure 4:

62 year-old female patient with A1AT grade moderate (FEV 43%); HP ^{3}He MRI (a) grade mild (25% non-ventilated lung) and CT (b) grade severe (65% diseased lung). This disparity suggests that ventilation is still taking place in extensive morphologically destroyed lung.





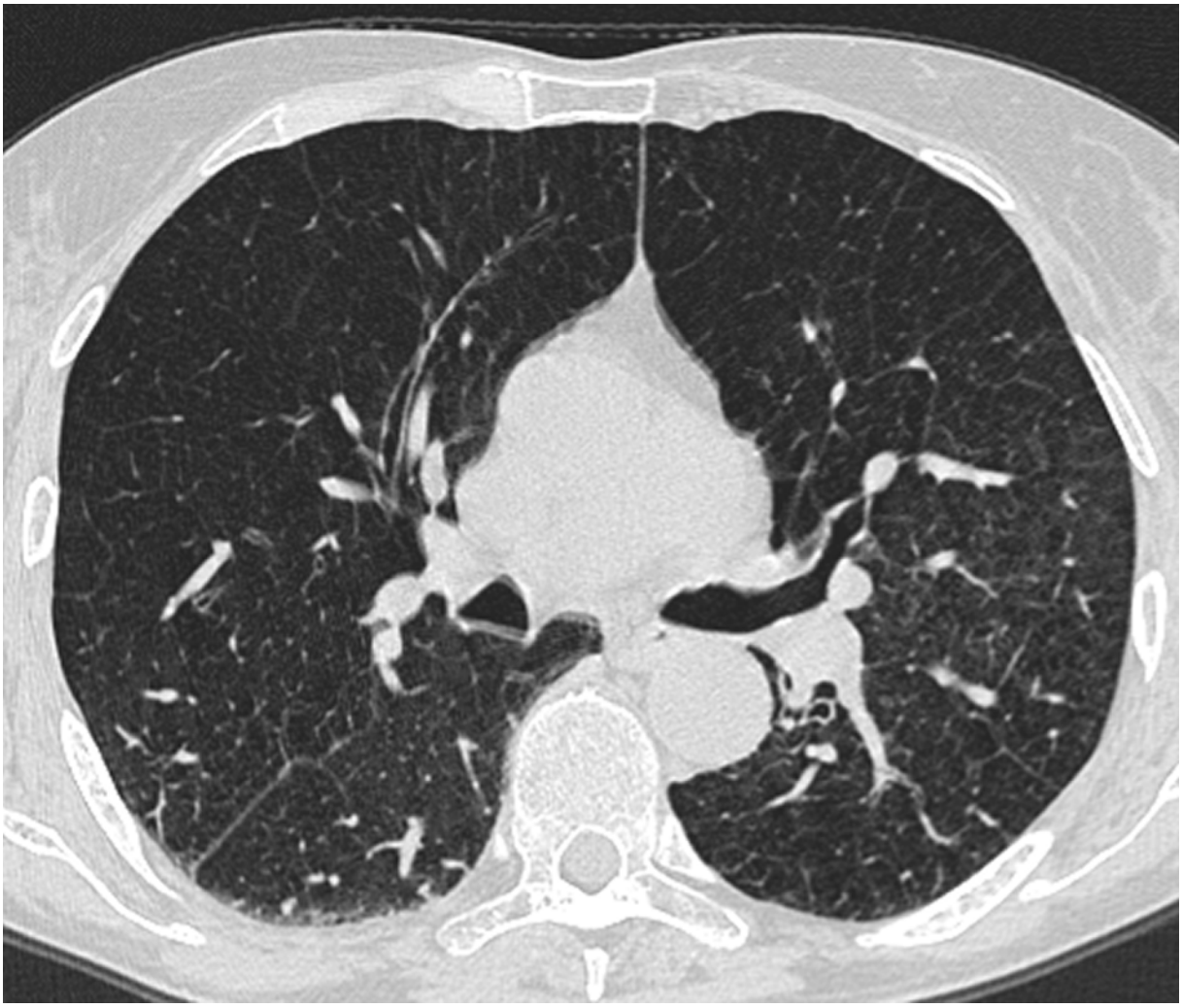
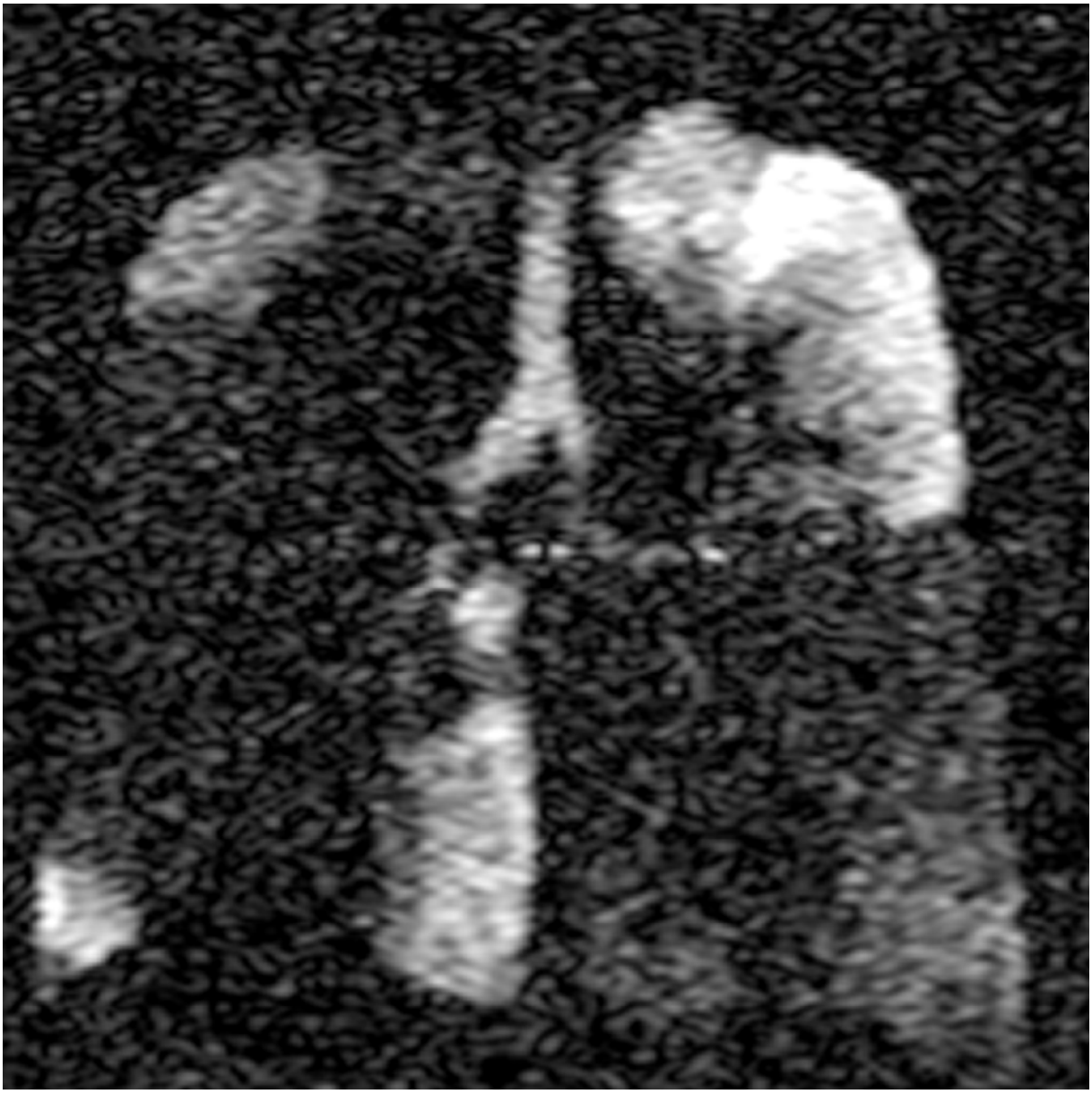


Figure 5:

65 year-old male patient with COPD grade severe (FEV 23%); HP ^{3}He MRI (a) grade moderate (45% non-ventilated lung) and CT (b) grade moderate (50% diseased lung).



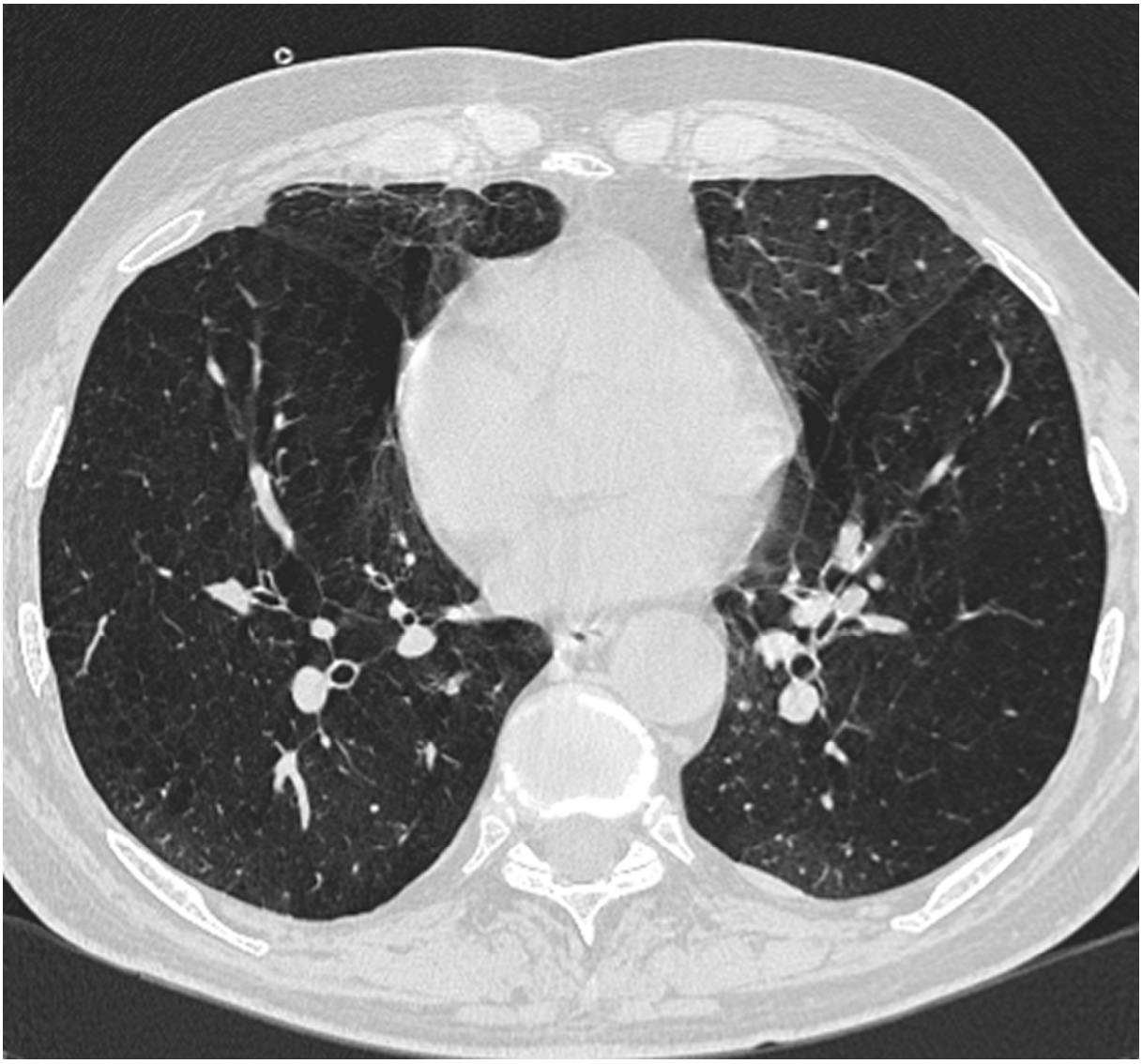
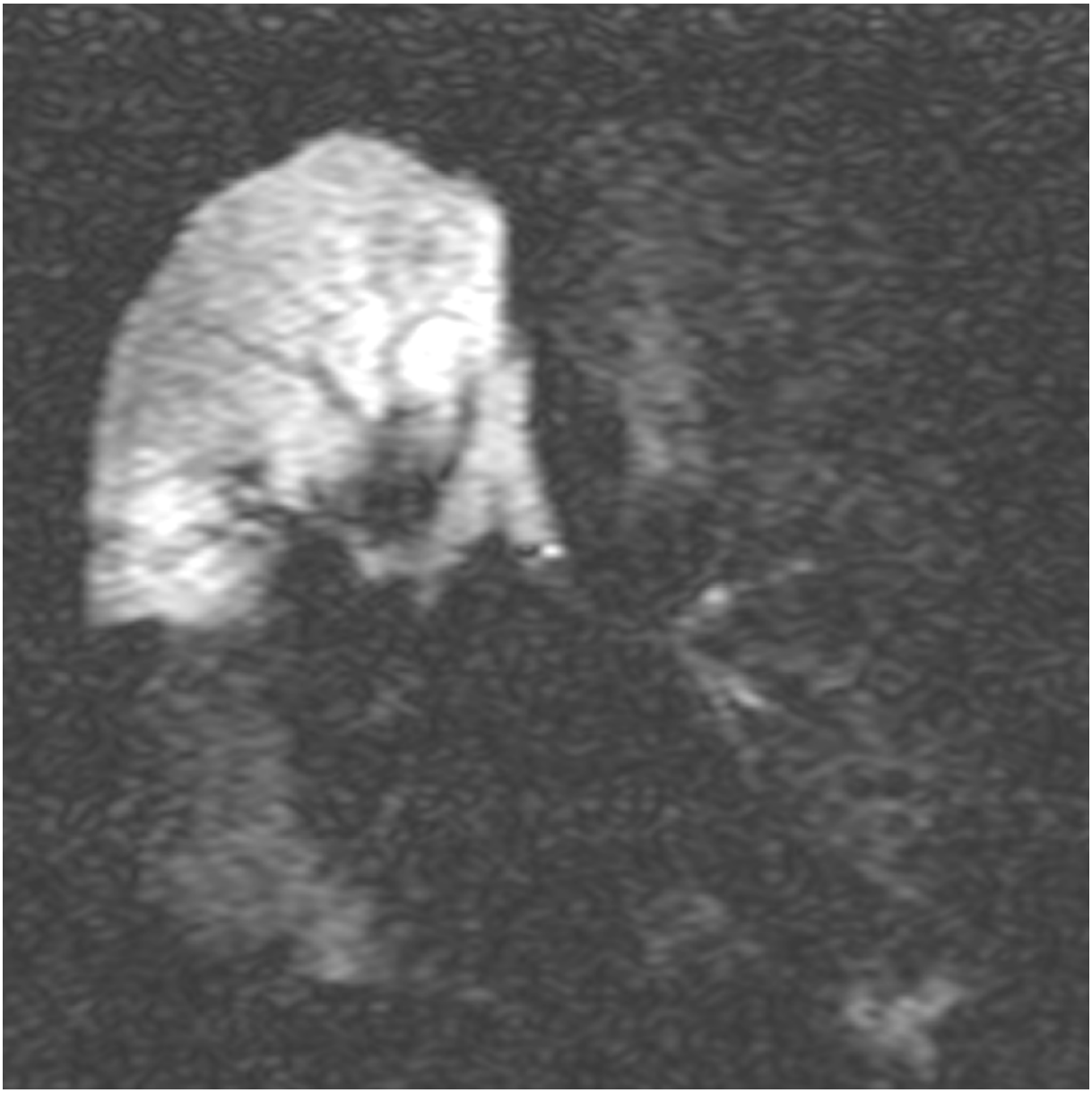


Figure 6:

50 year-old male patient with COPD grade severe (FEV 22%); HP ^{3}He MRI (a) grade moderate (60% non-ventilated lung) and CT (b) grade severe (80% diseased lung. In fact, for ^{3}He MRI the threshold for severe was >60%, so these findings (although statistically compatible with disagreement, clinically these findings are in agreement).



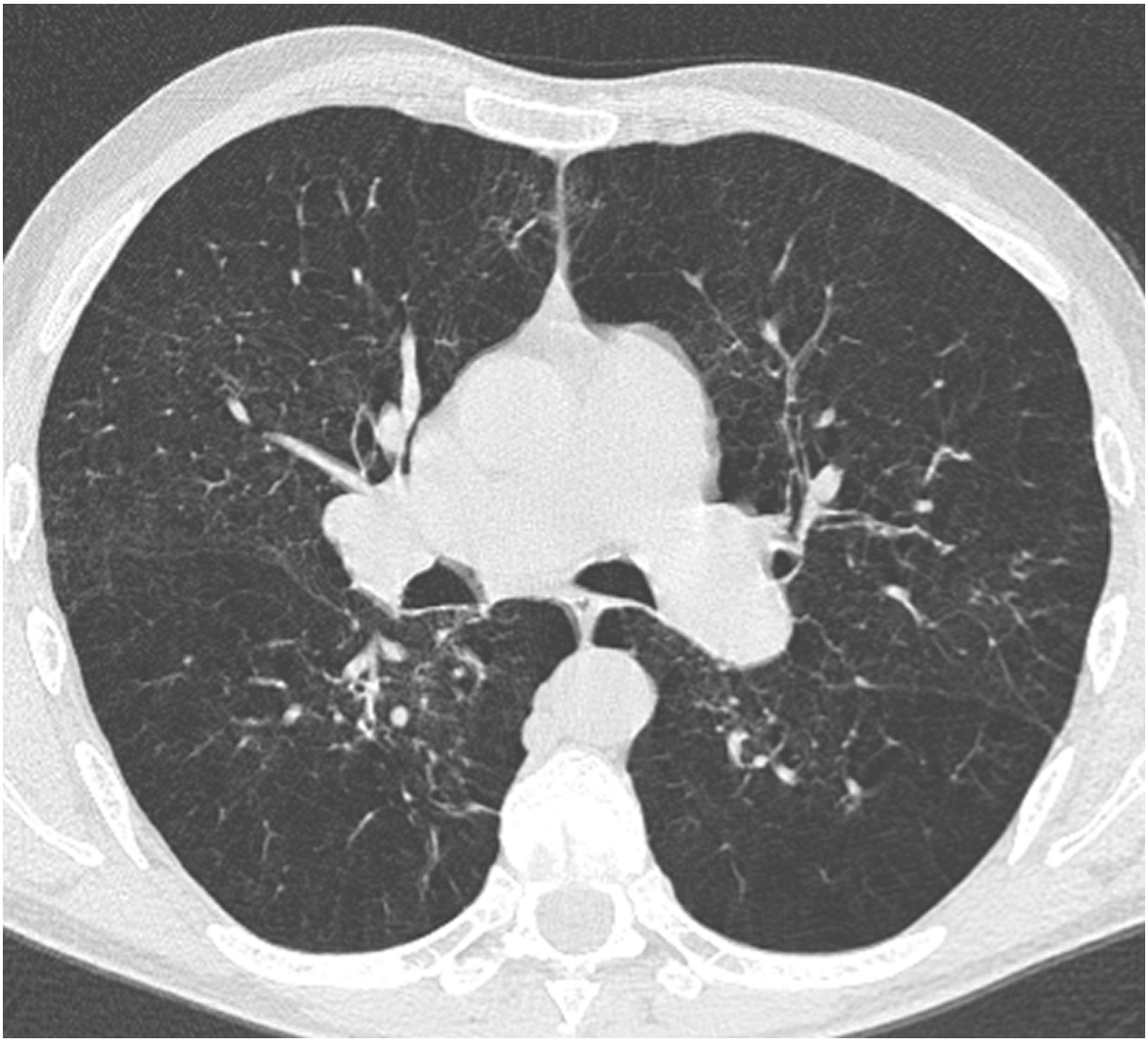


Figure 7:

Box whisker plots for the distributions of the apparent diffusion coefficient (ADC) and the mean lung density (MLD), respectively, stratified for the study subjects' health status; horizontals indicate medians and quartiles, verticals indicate maximum and minimum observations, circles indicate statistical outliers

Figure 7a

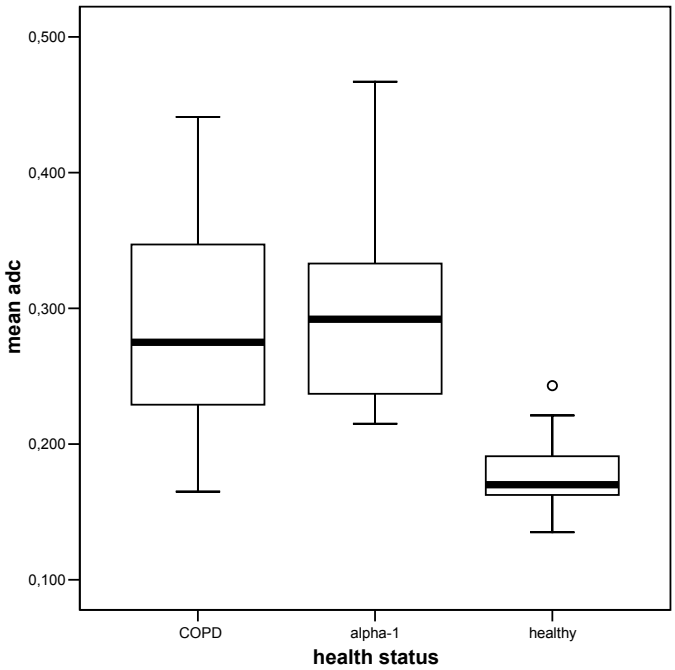
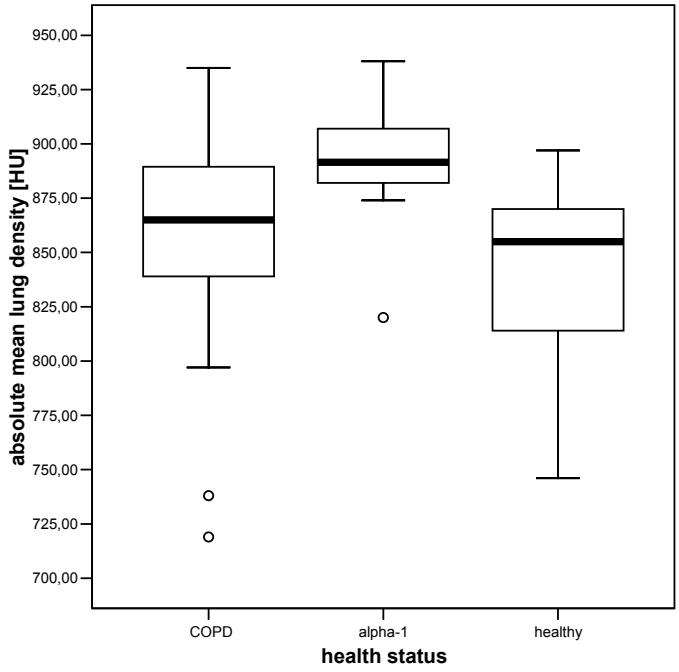

0,500
0,400
0,300
0,200
0,100
mean adc
COPD
alpha-1
healthy
health status

Figure 7b


950,00
925,00
900,00
875,00
850,00
825,00
800,00
775,00
750,00
725,00
700,00
absolute mean lung density [HU]
COPD
alpha-1
healthy
health status