Imaging biomarkers of lung ventilation in interstitial lung disease from $^{129}$Xe and oxygen enhanced $^1$H MRI

Marta Tibiletti $^a$, James A. Eaden $^b$, Josephine H. Naish $^{a,c}$, Paul J.C. Hughes $^b$, John C. Waterton $^{a,d}$, Matthew J. Heaton $^a$, Nazia Chaudhuri $^e$, Sarah Skeoch $^i$, Ian N. Bruce $^g,h$, Stephen Bianchi $^i$, Jim M. Wild $^{b,j}$, Geoff J.M. Parker $^{a,k,*}$

$^a$ Bioxydyn Limited, Rutherford House, Manchester Science Park, Manchester M15 6SZ, United Kingdom
$^b$ MCMR, Manchester University NHS Foundation Trust, Wythenshawe, Manchester, UK
$^c$ Centre for Imaging Sciences, University of Manchester, Manchester, UK
$^d$ Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London, UK
$^e$ MCR, University of Sheffield MRI Unit, Department of Infection, Immunity and Cardiovascular Disease, The University of Sheffield, Sheffield, UK
$^f$ Insigneo Institute for in silico medicine, Sheffield, UK
$^g$ Royal National Hospital for Rheumatic Diseases, Royal United Hospitals Bath NHS Foundation Trust, Bath, UK
$^h$ Centre for Musculoskeletal Research, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
$^i$ Centre for Imaging Sciences, University of Manchester, Manchester, UK
$^j$ MRC Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK
$^k$ Academic Directorate of Respiratory Medicine, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

$^*$ Corresponding author at: Centre for Medical Image Computing, 1st Floor 90 High Holborn, London WC1V 6LJ, UK.
E-mail address: geoff.parker@ucl.ac.uk (G.J.M. Parker).

Purpose: To compare imaging biomarkers from hyperpolarised $^{129}$Xe ventilation MRI and dynamic oxygen-enhanced MRI (OE-MRI) with standard pulmonary function tests (PFT) in interstitial lung disease (ILD) patients. To evaluate if biomarkers can separate ILD subtypes and detect early signs of disease resolution or progression.

Population: Forty-one ILD (fourteen idiopathic pulmonary fibrosis (IPF), eleven hypersensitivity pneumonitis (HP), eleven drug-induced ILD (DI-ILD), five connective tissue disease related-ILD (CTD-ILD)) patients and ten healthy volunteers imaged at visit 1. Thirty-four ILD patients completed visit 2 (eleven IPF, eight HP, ten DI-ILD, five CTD-ILD) after 6 or 26 weeks.

Field strength/sequence: MRI was performed at 1.5 T, including inversion recovery T1 mapping, dynamic MRI acquisition with varying oxygen levels, and hyperpolarised $^{129}$Xe ventilation MRI. Subjects underwent standard spirometry and gas transfer testing.

Assessment: Five $^1$H MRI and two $^{129}$Xe MRI ventilation metrics were compared with spirometry and gas transfer measurements.

Statistical test: To evaluate differences at visit 1 among subgroups: ANOVA or Kruskal-Wallis rank tests with correction for multiple comparisons. To assess the relationships between imaging biomarkers, PFT, age and gender, at visit 1 and for the change between visit 1 and 2: Pearson correlations and multilinear regression models.

Results: The global PFT tests could not distinguish ILD subtypes. Percentage ventilated volumes were lower in ILD patients than in HVs when measured with $^{129}$Xe MRI (HV 97.4 $\pm$ 6.5, DI-ILD: 91.0 $\pm$ 4.8, CTD-ILD: 91.1 $\pm$ 7.4, HP 92.6 $\pm$ 4.0 = 0.013, IPF 88.1 $\pm$ 6.5 $p < 0.001$), but not with OE-MRI. $^{129}$Xe reported more heterogeneous ventilation in DI-ILD and IPF than in HV, and OE-MRI reported more heterogeneous ventilation in DI-ILD and IPF than in HP or CTD-ILD. The longitudinal changes reported by the imaging biomarkers did not correlate with the PFT changes between visits.

Data conclusion: Neither $^{129}$Xe ventilation nor OE-MRI biomarkers investigated in this study were able to differentiate between ILD subtypes, suggesting that ventilation-only biomarkers are not indicated for this task. Limited but progressive loss of ventilated volume as measured by $^{129}$Xe-MRI may be present as the biomarker of...
1. Introduction

Interstitial lung diseases (ILD) constitute a heterogeneous group of conditions exhibiting inflammation and scarring of the lung parenchyma. Pathological changes are spatially heterogeneous with varying degrees of acute inflammation and fibrosis. The ILDs have varying aetiologies and natural history. Typically, idiopathic pulmonary fibrosis (IPF), the most common ILD, has a chronic progressive phenotype, whereas other sub-types such as drug-induced ILD (DI-ILD) [1] and hypersensitivity pneumonitis (HP) [2] may reverse following withdrawal of the trigger. Diagnosis of ILDs remains a challenge, with a requirement for a multi-disciplinary assessment combining clinical history, immune profiling, lung physiology, computed tomography (CT), and lung biopsy [3]. It is important to accurately classify a subject’s ILD subtype as this has an impact on the prognosis as well as the choice of the most effective treatment for the patient, e.g., antifibrotics in IPF [4] and immunosuppressants for other subtypes [1]. Thus, there is a need for improved biomarkers for precise diagnosis and monitoring of disease progression and treatment efficacy. Pulmonary function tests (PFTs) lack disease specificity as they measure the global function of the lungs only [5] and cannot interrogate regional change in ILD, unlike imaging biomarkers, which provide regional information [6]. Numerous observational studies have reported cohort predictors of ILD progression and/or mortality [7,8]; however, accurate prognosis for individual ILD patients remains a challenge. Most imaging biomarker studies in ILD have focused on IPF [9] in small numbers of subjects without independent validation [3].

Where repeated assessments are required, it is desirable to avoid ionising radiation, and risks from imaging contrast agents should be minimised. MRI biomarkers therefore are of particular interest, particularly when benign inhaled gases such as oxygen or hyperpolarised 129-xenon (129Xe) are used to provide additional structural and physiological information [10–15]. Hyperpolarised 129Xe MRI exploits the signal enhancement available following spin exchange optical pumping to allow for the direct visualisation of inhaled gases and ventilation at high resolution [16]. In this work, only 125Xe ventilation MRI is considered, and more complex techniques based on spectroscopy of dissolved 128Xe [9,16] or diffusion-weighted MRI [15] were not included.

The technology to produce and visualise hyperpolarised gases in clinical settings is currently limited to a few specialised centres [17], as hyperpolarised 129Xe is categorised as an investigational medicinal product and the expense of the added equipment and personnel required limits its widespread use. For this reason, alternative methodologies to image ventilation are of interest. One candidate is oxygen enhanced MRI (OE-MRI), which exploits the effect of molecular oxygen on lung tissue water in conventional proton MRI. Pure oxygen and medical air are widely available in hospital settings, and their delivery can be reliably achieved with standard medical equipment [18].

OE-MRI permits quantification of change in concentration of dissolved oxygen in lung tissues induced by inhaling changed concentrations of O2 [19], and has previously been deployed in other lung diseases [20,21]. OE-MRI is usually paired with a measurement of native T1 in the parenchyma, itself a promising imaging biomarker of focal lung disease [22,23]. There is still little published information on the performance of such MR biomarkers in ILD, their temporal evolution, and correlations between the different MR biomarkers and conventional pulmonary function tests.

In this study we aimed: firstly, to compare imaging biomarkers derived from 3H-T1 mapping, dynamic OE-MRI and hyperpolarised 129Xe ventilation MRI with standard lung physiological measurements in ILD patients in comparison to healthy volunteers (HV); secondly, to compare these MRI biomarkers between ILD subtypes; and thirdly, to assess longitudinal changes in these MRI biomarkers.

2. Materials and methods

2.1. Participants

The study was carried out as part of a programme to validate imaging biomarkers of drug safety [24]. It was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the local research ethics committee (United Kingdom North West -Preston Research Ethics Committee, REC Ref 17/NN/0631, IRAS number: 232495). The study investigated several imaging biomarkers intended to probe lung morphology, perfusion, and ventilation. This report includes only the subgroup of patients in whom OE-MRI and 129Xe ventilation MRI was performed.

Patients with a diagnosis of ILD were recruited and assigned to one of four ILD subtypes: suspected DI-ILD, connective tissue disease related-ILD (CTD-ILD), IPF or HP.

The diagnosis of the ILD subtype was established in ILD multidisciplinary team (MDT) meetings involving respiratory physicians, thoracic radiologists, and pathologists, the gold standard for the diagnosis of ILD since the publication of the ATS/ERS IIP classification in 2002 [25]. Potential subjects were identified by respiratory physicians during ILD MDT meetings where patients’ cases were discussed as part of routine clinical care. Current diagnostic investigations in ILD mainly consist of HRCT and PFTs [25]. It should be noted that often a definitive diagnosis cannot be achieved by the MDT, but instead a “working diagnosis” of high probability can be reached by combining the key information available to increase or decrease the diagnostic probability of a specific ILD subtype [26]. As this was an observational trial, recruited patients followed the usual standard of care, but treatment regime was noted.

Exclusion criteria were: significant pre-existing cardio-pulmonary disease, radiotherapy to the lung fields within 6 months, any features of any malignancy involving the lungs, evidence of lower respiratory infection, estimated survival <6 months and any contraindication for MR imaging.

Patients were recruited, gave written informed consent, and underwent a clinical assessment, lung function testing and MRI at visit 1. Patients were recalled for a follow up visit after 6 weeks (if diagnosed with suspected DI-ILD or HP) or 6 months (if diagnosed with IPF or CTD-ILD). The follow up schedules were tailored to enable the capture of change at the most appropriate time points for each disease group, as IPF patients typically show a slower progressive decline while DI-ILD and HP often either rapidly declines or improves with treatment. Additionally, ten healthy volunteers (HV) were also recruited under ethical approval provided by the UK national research ethics committee (REC Ref 12/NE/0355), with all volunteers giving written informed consent to undergo pulmonary function testing and MRI at baseline only.

2.2. Pulmonary function tests

Patients underwent pulmonary function testing prior to MRI scanning, performed to international standards [27] under the supervision of a trained respiratory physiologist. Spirometry was performed to assess the forced expiratory volume in 1 s (FEV1) and forced vital capacity (FVC), Gas transfer testing assessed the transfer factor (TLCO) and coefficient (KCO2) of the lungs for carbon monoxide [28,29]. All values are reported as % predicted using the Global Lung Function Initiative (GLI)
Table 1

Imaging parameters for $^{129}$Xe and OE acquisitions. SPGR: spoiled gradient echo, SSFP: steady state free precession, IR-TFE: inversion recovery turbo spin echo, TE: echo time, TR: repetition time.

<table>
<thead>
<tr>
<th>Metric</th>
<th>$^1$H (SPGR)</th>
<th>$^{129}$Xe ventilation (SSFP)</th>
<th>$T_1$ Mapping (IR-TFE)</th>
<th>Dynamic OE (IR-TFE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition matrix</td>
<td>100 × 100</td>
<td>100 × 80</td>
<td>96 × 96</td>
<td>96 × 96</td>
</tr>
<tr>
<td>Pixel size (mm$^2$)</td>
<td>4 × 4</td>
<td>4 × 4</td>
<td>4.2 × 4.2</td>
<td>4.2 × 4.2</td>
</tr>
<tr>
<td>Slice thickness (mm)</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Flip angle (degrees)</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>TE/TR (ms)</td>
<td>0.6/1.9</td>
<td>2.2/6.7</td>
<td>0.4/1.5</td>
<td>0.4/1.5</td>
</tr>
<tr>
<td>BW (kHz)</td>
<td>±83.3</td>
<td>±8.0</td>
<td>±31.25</td>
<td>±31.25</td>
</tr>
</tbody>
</table>

2012 reference eqs. [30].

2.3. Imaging

All imaging was carried out in the coronal plane on a 1.5 T whole body system (GE HDx, GE Healthcare, Milwaukee, WI): Table 1 provides details on the sequences used.

2.3.1. Hyperpolarised $^{129}$Xe imaging

Hyperpolarised ventilation $^{129}$Xe imaging was performed using a flexible quadrature transmit/receive quadrature coil (Clinical MR Solutions, Brookfield, Wisconsin, USA). $^{129}$Xe was polarised under regulatory licence to $\sim$30% using an in-house spin exchange optical pumping polariser capable of generating 500 ml doses in $<15$ min [31]. $^{129}$Xe images were acquired at a breath-hold of functional residual capacity plus 1 l of gas mixture (FRC + 1 l, 500 ml $^{129}$Xe, 500 ml N$_2$) with a 3D balanced steady-state free precession (bSSFP) sequence with a 10 mm slice thickness and pixel size of 4 mm $\times$ 4 mm. Prior to the acquisition of the $^{129}$Xe images, a $^1$H structural 3D spoiled gradient echo (SPGR) image was acquired at FRC + 1 l utilising a 1 l bag of N$_2$. The $^1$H structural image had the same in-plane spatial resolution, but a slice thickness of 5 mm compared to 10 mm for the $^{129}$Xe image to allow for an affine co-registration and the estimation of the percentage lung ventilated volume (Xe-VVF) [32]. To ensure all subjects were able to complete the breathing manoeuvre, appropriate training took place prior to imaging.

2.3.2. Inversion recovery $T_1$ measurement and dynamic oxygen-enhanced imaging

Following completion of $^{129}$Xe imaging, subjects were repositioned in an 8-element $^1$H chest receiver coil and were fitted with a disposable non-rebreathing mask to allow for medical air and oxygen delivery. A free-breathing protocol based on an inversion-prepared centric ordered single shot 3D turbo field echo (IR-TFE) sequence was used [33] with 10 mm slice thickness and 4.2 mm $\times$ 4.2 mm spatial resolution. A baseline $T_1$ map was calculated from 6 acquisitions with variable inversion time (TI) (40, 100, 300, 1100, 2000 and 5000 ms). Five volumes were acquired for each TI, to capture different stages of the respiratory cycle. A dynamic OE acquisition followed, lasting 15 min (TI = 1100 ms) with a temporal resolution of 10 s, during which gas was delivered at 15 l/min and switched from medical air to 100% O$_2$ at minute 2, and back to air at minute 10 [34].

2.4. Image analysis

2.4.1. $^{129}$Xe image analysis

$^1$H structural images were co-registered to the same spatial domain as the $^{129}$Xe ventilation image and segmented semi-automatically using spatial fuzzy c-means thresholding and manual editing [31]. Xe-VVF was calculated by dividing the ventilated volume (from the ventilation image) by the thoracic cavity volume (from the $^1$H structural image). Additionally, the median and interquartile range (IQR) of the coefficient of variation (CV) of ventilated signal intensity was calculated [35,36], with the IQR of the whole lung CV being referred to as the xenon ventilation heterogeneity index (Xe-VH). Briefly, for CV calculation the $^{129}$Xe images were subsampled in-plane by 50% and a 3 $\times$ 3 sliding window used to calculate voxel-wise CV, incorporating only voxels identified as ventilated (from the ventilated volume mask).

2.4.2. Inversion recovery $T_1$ and dynamic oxygen-enhanced imaging analysis

All OE-MRI images were co-registered using ANTs [37] to a reference image representing expiration. Baseline $T_1$ was calculated by fitting the inversion recovery data as previously described [34].

The lung cavity was semi-automatically segmented from the reference image using a region growing algorithm from manually defined points. Lung volume changes during the dynamic acquisition were estimated from the mask and the deformation field extracted from the registration. The registration was assessed by extracting the apparent diaphragm displacement from the dynamic images after motion-correction: any frame presenting an apparent post-correction displacement $>1$ pixel was excluded from further analysis.

The dynamic signal at each pixel in the lung parenchyma was modelled by sum of two signals: (1) a component dependent on lung volume change, describing the local variation in local proton density during the respiratory cycle and (2) a piecewise mono-exponential recovery function in the time domain, modelling the local increase in $T_1$-weighted MR signal due to an increase in concentration of dissolved molecular oxygen after gas switching.

To account for the first element, a first-degree polynomial fit between the log-transformed pixelwise MR signal and the log-transformed whole lung volume changes was calculated and subtracted from the original signal. The oxygen enhancement was then derived by the change in signal intensity between the median values on medical air and the median values of the last ten frames on 100% O$_2$. This signal change was then converted to the change in $R_1$ [38], and then into a change of partial pressure of oxygen ($\Delta pO_2$) using the longitudinal relaxivity in water ($R_1$) of $2.49 \times 10^{-4}$ mmHg [39].

In line with the $^{129}$Xe analysis, the median value of the coefficient of variation of the $\Delta pO_2$ map (median CV $\Delta pO_2$) was extracted by a 3 $\times$ 3 2D kernel. The interquartile range of this map was also calculated, as the oxygen enhanced ventilation heterogeneity index (OE-VH).

The oxygen wash in time ($\tau_{wp}$) was estimated by fitting the signal with a piecewise mono-exponential curve. When the Akaike information criterion (AIC) favoured the latter fitting over a constant function, the pixel was considered as ventilated, which allowed the calculation of the oxygen enhanced ventilated volume fraction (OE-VVF). The OE-VVF was applied to the $\tau_{wp}$ map as a mask to exclude pixels with no detectable oxygen enhancement.

2.5. Statistics

The five $^1$H MRI biomarkers ($T_1$, $\Delta pO_2$, $\tau_{wp}$, OE-VH$^1$ and OE-VVF$^1$) and the two $^{129}$Xe ventilation biomarkers (Xe-VVF and Xe-VH), were compared with the spirometry (FEV$_1$/FVC), and the gas transfer biomarker TLCO$.%$. To evaluate differences in biomarkers at visit 1 among the ILD subgroups and the healthy volunteers, ANOVA or Kruskal-Wallis rank tests followed by post hoc test with Bonferroni or Dunn correction for multiple comparisons were carried out (depending on the result of the Shapiro-Wilk normality test). The level of significance was set at $p = 0.05$ for these tests after multiple correction.

Also, the considered population data at visit 1 was divided into three groups: healthy volunteers, ILD subject with TLCO$\% < 75$ (High TLCO$\%$ and ILD subject with TLCO$\% \leq 75$ (Low TLCO$\%$). A similar analysis to the one just described was carried out to determine if any
imaging biomarker could separate HV from both High TLCO% and Low TLCO% groups.

To assess the relationships between imaging biomarkers, Pearson correlations were carried out. Pearson correlations were also carried out between each of the imaging biomarkers and PFTs and age/gender, both at visit 1 and (separately) for the between visits change. The level of significance was set at $p = 0.01$ for these tests. When at least one significant correlation was found, the best multilinear regression model was identified as having had the imaging biomarker as an independent parameter, the PFTs and age as a continuous covariate and sex and disease status (HV, IPF, DI-ILD, CTD-ILD, HP) as categorical variables, using a stepwise method guided by a decreasing AIC. The results of the multivariate models are presented as the coefficient $\beta$ and its 95% confidence interval and $R^2$, the percentage of variation in the response that is explained by the model.

To assess the change of PFT between visits in the whole population, and in each of the ILD subgroups, a one sample t-test or Wilcoxon signed rank test was carried out. The level of significance was set at $p = 0.01$ for these tests. No correction for multiple comparisons was applied.

Moreover, subjects were divided in two groups depending on whether they received a pharmacological treatment for ILD during the study. The change in biomarker between visits was compared between the two groups with a two-tailed t-test or a Wilcoxon signed rank test depending on the result of the Shapiro normality test. The level of significance was set at $p = 0.01$ for these tests. No correction for multiple comparisons was applied. All statistical analysis were run using Python libraries SciPy (version 1.6.0) and statsmodels (version 0.12.0).

![Fig. 1. Comparison of obtained images from $^{129}$Xe ventilation (first row, with CT) and T1 and OE-MRI (second row) in a single slice for a subject diagnosed with IPF (FEV1% 85.8, FVC% 66.3, TLCO% 26.6, KCO% 44).](image1)

![Fig. 2. Front to back slice by slice comparison of $^{129}$Xe images (first row) and OE-MRI $\Delta pO_2$ enhancement (second row) and oxygen wash in rate (third row) a subject affected by Hypersensitive Pneumonitis. Ventilation appears to be fairly homogeneous with both modalities. The main ventilation defect visible on the $^{129}$Xe images is located in the upper right lobe and indicated by an asterisk. The same area presents normal oxygen enhancement levels but high oxygen wash-in rate. Apparent artefacts are visible close to the hearth in the left lung in OE-MRI images [19].](image2)
3. Results

Fig. 1 shows one representative slice from all considered biomarker maps, $^{129}$Xe biomarker, CT, and OE-MRI for an IPF subject. Figs. 2 and 3 show a comparison of anterior-posterior slices from the $^{129}$Xe volumetric acquisition and OE-MRI enhancement map and wash-in rate of a subject with HP and a patient with IPF. A clear gradient in $^{129}$Xe spin density front-to-back is clearly visible in $^{129}$Xe ventilation images. This gradient seems reversed in $\Delta pO_2$ images, with posterior images enhancing less than anterior ones. In Fig. 2 ventilation is mostly uniform, but a ventilation defect visible on $^{129}$Xe images is visible in the upper right lung and corresponds to normal $\Delta pO_2$ enhancement but high $t_{up}$. Apparent artefacts are visible in the left lung, close to the heart, in OE-MRI images. In Fig. 3, significant differences in contrast in between modalities are clearly visible, but some commonality in ventilation defects are also evident (arrow, asterisk and plus signs).

3.1. Population characteristics

Fig. 4 summarises in a flow chart the patient recruitment in the study. Forty-one ILD patients were recruited (14 IPF, 11 HP, 11 DI-ILD and 5 CTD-ILD) and successfully imaged at visit 1. Thirty-six complete datasets of ILD patients imaged at visit 1 were analysed (12 IPF, 10 HP, 11 DI-ILD and 3 CTD-ILD). Five subjects were not analysed; one subject due to the dynamic OE-MRI protocol not being fully performed and four subjects OE-MRI datasets were excluded due to absent or weak $O_2$ enhancement in the blood pool as measured in the aorta, which indicated issues with the gas delivery during scanning. Thirty-four ILD patients attended and completed visit 2 (11 IPF, 8 HP, 10 DIILD, 5 CTD-ILD). Of the seven patients who failed to attend, three subjects died between visit 1 and visit 2, two withdrew from the study, one was lost at follow up and one could not be scanned in the appropriate time window due to MR hardware issues. One of the thirty-four acquisitions could not be analysed due to inconsistent field of view prescription among the OE acquisitions. No incidental findings were reported in the study population.
A total of twenty-nine ILD patients (nine IPF, ten HP, seven CTD-ILD, three CTD-ILD) completed both visits and had analysable OE-MRI datasets. One HV was excluded from the analysis due to a radiological incidental finding, and the remaining nine HV were analysed. Table 2 summarises biomarker results obtained at visit 1 and the change between visits (visit 2 – visit 1).

Regarding pharmacological treatment, in the IPF group, four subjects were treated with an antifibrotic (Nintedanib), while the remaining subjects were not under pharmacological treatment. Among the subjects diagnosed with HP, all were treated with a corticosteroid (prednisone); in addition to this, one HP patient also received Azathioprine (AZA) and four patients also received Mycophenolate Mofetil (MMF). In the CTD-ILD group, one subject received no pharmacological treatment during the study, two were on a corticosteroid plus AZA and two on a corticosteroid plus MMF. Among subjects diagnosed with DI-ILD, six received a corticosteroid, while five did not receive drugs for the condition as their management plan involved withdrawal of the causative drug only. Starting, stopping and length of treatment also varied widely during the study.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>DHILD</th>
<th>IPF</th>
<th>HP</th>
<th>CTD-ILD</th>
<th>HV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>11</td>
<td>14</td>
<td>56</td>
<td>45</td>
<td>9</td>
</tr>
<tr>
<td>Age [y]</td>
<td>96.4 ± 7.8</td>
<td>125 ± 17.5</td>
<td>10.9 ± 17.4</td>
<td>11.0 ± 10.7</td>
<td>11.7 ± 9.4</td>
</tr>
<tr>
<td>FEV1 % v1</td>
<td>84.2 ± 93.5</td>
<td>74.4 ± 89.0</td>
<td>95.2 ± 56.9</td>
<td>54.1 ± 48.6</td>
<td>54.6 ± 48.1</td>
</tr>
<tr>
<td>Xe-VVF v1</td>
<td>9.3 ± 7.3</td>
<td>55.5 ± 37.3</td>
<td>5.5 ± 5.5</td>
<td>5.5 ± 5.5</td>
<td>5.5 ± 5.5</td>
</tr>
<tr>
<td>Xe-VVF (v2–v1)</td>
<td>24.4 ± 25.2</td>
<td>12.7 ± 9.6</td>
<td>9.6 ± 6.6</td>
<td>9.6 ± 6.6</td>
<td>9.6 ± 6.6</td>
</tr>
<tr>
<td>Xe-VVF (v2–v1)</td>
<td>24.4 ± 25.2</td>
<td>12.7 ± 9.6</td>
<td>9.6 ± 6.6</td>
<td>9.6 ± 6.6</td>
<td>9.6 ± 6.6</td>
</tr>
</tbody>
</table>

### 3.2. PFTs and imaging biomarkers at visit 1

No difference was found in FEV1% and FVC% between HV and any ILD subgroups. HV had higher TLCO% (p < 0.001) than all ILD groups, while HV had statistically higher KCO% than IPF (p < 0.0001), DI-ILD (p = 0.01) and CTD-ILD (p = 0.01), but not HP (p = 0.12). There were no significant differences between ILD subgroups for these biomarkers (Fig. 5). The subgroup considered differed in age: HV (49.4 ± 17.4 y) were significantly younger than the IPF (IPF: 71.9 ± 7.18 y, p = 0.004). IPF patients were also older than CTD-ILD patients (58.5 ± 10.9 y, p = 0.04) (Table 2).

Fig. 6 presents the boxplot of the considered 129Xe biomarkers and Fig. 7 presents the boxplot of the considered OE-MRI biomarkers, divided by ILD subgroups. Regarding 129Xe biomarkers, the ventilation volume fraction was lower in all ILD groups than in HV (Xe-VVF: HV mean ± std. 97.4 ± 2.6, CTD-ILD: 91.0 ± 4.8, p = 0.017, DI-ILD 90.1 ± 7.4 p = 0.003, HP 92.6 ± 3.8 p = 0.013, HP 88.1 ± 6.3 p < 0.001), but this was not replicated with the OE ventilation volume fraction. As for PFTs, when averaged across the lung, the OE-MRI biomarkers generally failed to distinguish the ILD subgroups, with the exception that ApO2 was higher in CTD-ILD than in IPF (p = 0.016).

There were however significant differences in ventilation heterogeneity between ILD groups. Xe-VH1 was lower in HV than in IPF (p < 0.001), Xe-VH1 was also lower in the CTD-ILD group than in IPF (p < 0.001), and lower in HP than in IPF (p = 0.042).

If visit 1 subjects are divided between HV, High TLCO% and Low TLCO% groups, Xe-VVF was higher in HV than in both ILD groups (HV 97.38 ± 2.64, High TLCO%: 94.11 ± 5.317, Low TLCO%: 90.25 ± 6.38%, HV vs High p = 0.002, HV vs Low, p < 0.001).

Xe-VH was significantly lower in HV (0.095 ± 0.013) than in High TLCO% (0.123 ± 0.02, p = 0.006), but not significantly different than in the High TLCO% (0.105 ± 0.021).

Similarly, T1i is higher in HV (1180.7 ± 95.5 ms) than in Low TLCO% (1103.4 ± 63.2 ms, p = 0.02), but not in High TLCO% (1157.68 ± 83 ms). T2p was lower in HV (38.77 ± 12.68 s) than in High TLCO% (49.32 ± 20.18 s, p = 0.03), but not significantly different than in Low TLCO% (48.8 ± 24.4 s).

Table 3 shows the Pearson correlation coefficients between imaging biomarkers at visit 1. No significant correlation was found between Xe-VVF and OE-VVF, nor between Xe-VH and OE-VH. There was a weak but significant correlation between T1i and VVF (R = –0.38, p = 0.009).

Multiplicative regression models applied to the whole considered population demonstrated that age (β = –0.2; 95%CI = −0.33, −0.11; p < 0.001) and TLCO (β = 0.07; 95%CI = 0.053, 0.136; p = 0.035) were significantly correlated with Xe-VVF (R2 = 0.39).

Age (β = 0.0007; 95%CI = −0.04, 0.05; p = 0.001), TLCO (β = −0.003; 95%CI = 0.034, 0.037; p = 0.013) and gender (β = 0.019; 95%CI = 0.009, 0.031; p = 0.001) were significantly correlated with Xe-VH (R2 = 0.53).

Age (β = −2.32; 95%CI = −3.85, −0.79; p = 0.004), FVC% (β = 0.97; 95%CI = 0.13, 1.81; p = 0.025), and IPF diagnosis (β = 33.2; 95%CI =...
14.53, 80.99; \( p = 0.168 \) were correlated with \( T_1 \) \( (R^2 = 0.24) \). No other associations were observed. The full table of Pearson correlation coefficients is available in the Supplementary Materials (Table 1).

3.3. Changes between visits

Boxplots representing the longitudinal changes in the considered biomarkers are visualised in the Supplementary Materials Figs. 1–3. Regarding longitudinal changes, no statistically significant changes were found in individual ILD subgroups or in the overall ILD population.

If the population is divided between subjects who received pharmacological treatment for ILD during the study, and subjects who did not, no significant difference in the biomarker changes were found between the two groups.

3.4. Multilinear models – imaging biomarkers (visit 2 – visit1)

Only age at visit 1 was a significant predictor of the change in VVF between visits \( (R^2 = 0.145; \beta = 0.2; \text{95\%CI} = 0.026, 0.44; \ p = 0.029) \). Multilinear models indicated that change of FEV1% was significantly correlated with the change in \( T_1 \) between visits \( (R^2 = 0.138, \beta = -0.47; \text{95\%CI} = -9.33, -0.157; \ p = 0.043) \) but no further associations with MRI biomarkers were observed. The full table of Pearson correlation coefficients is available in the Supplementary Materials (Table 2).

4. Discussion

Despite the complexity of the imaging employed in the study, only 2 subjects withdrew, demonstrating the feasibility of multi-sequence MR acquisition in ILD, also in patients with severely deteriorated lung function. Issues with oxygen delivery during scanning were detected – methods to identify when gas delivery is failing during scanning may be helpful in the future to improve data quality.

In this study, none of the global measurements (PFT, average \( T_1 \), or any of the global OE-MRI or \( ^{129} \text{Xe} \) ventilation biomarkers) differentiated between ILD subtypes, although biomarkers of focal variation (OE-VH and Xe-VH) may show differences in ventilation heterogeneity between ILD subtypes. The importance of accurately and quickly classifying a subject’s ILD subtype comes from the impact on the prognosis as well as the choice of the most effective treatment for the patient (e.g., immunosuppression versus antifibrotic therapy). The results from this study suggest that ventilation-only biomarkers are not suitable for this specific objective, nor to track the longitudinal changes. In fact, the longitudinal change in the imaging biomarkers did not correlate with PFT changes.
between the two timepoints. The significant correlation between \( \tau_{up} \) and Xe-VVF suggests that areas of the lung not reached by \( ^{129}\text{Xe} \) during the breath-hold experiment may be reached by oxygen at a slower rate during the significantly longer free breathing OE-MRI experiment. This would also be consistent with the observation that, unlike Xe-VVF, the fraction of lung apparently ventilated by oxygen (OE-VVF) did not differentiate between groups.

The only imaging biomarker that could separate the healthy volunteers and both the high and low TLCO% groups was Xe-VVF, which indicates that a limited but progressive loss of ventilated volume may occur as disease progresses.

The ILD patients recruited at baseline had on average near normal FVC% and FEV_1%, but similar low TLCO%\%; it is perhaps therefore unsurprising that functional imaging contrasts that are dominated by ventilation were unable to differentiate these groups [30]. Perhaps more surprisingly, as noted above, ventilated volume fractions did not correlate between OE-MRI and \( ^{129}\text{Xe-MRI} \). In addition to the differences in the time scales over which the measurements are performed, this lack of correlation is possibly because OE-MRI measures signal enhancement in the parenchyma and blood in the lungs following oxygen ventilation, and is therefore best considered as creating contrast weighted by both ventilation and perfusion, while ventilation \( ^{129}\text{Xe-MRI} \) directly measures ventilation in the airways and the alveoli; alternatively, it may simply reflect the relative variability and small numbers in this work.

The OE-VVF obtained in this work is relatively low also in healthy volunteers (80.8 ± 9.8%). Lack of enhancement is expected in bronchi and arteries in this imaging method, therefore reducing the achievable VVF to <100%. It is also possible that movement artefacts from breathing, the cardiac cycle and bulk movements may contribute to decreased OE-VVF.

Also, while both OE-MRI and \( ^{129}\text{Xe-MRI} \) suggested more heterogeneous ventilation in DI-ILD and IPF, ventilation heterogeneity index measurements failed to correlate between the two modalities, potentially for the same underlying reasons.

Correlation between Xe-VVF obtained with \( ^{3}\text{He} \) imaging and OE-VVF has been seen in a cystic fibrosis study [40] which did not investigate oxygen wash-in time. The lung pathology typical of CF is characterised by airway obstruction, leading to markedly decreased FEV_1 and FVC, markedly different from the fibrotic pathology present in ILD. Further studies are necessary to investigate the apparent distribution of different gases as visualised by MR imaging in an array of pulmonary diseases.

It is well known that \( ^{129}\text{Xe} \) ventilation biomarkers correlate strongly with age [41]. When age is considered in a multilinear model, Xe-VVF and VH_I are significantly correlated with TLCO\% at visit 1. This result also indicates that reduction in ventilation volume and increase in ventilation heterogeneity in subjects with higher thickening of the alveolar-capillary membrane.

Sex is also independently correlated with VH_I with female subjects presenting more uniformly ventilated lungs than male. Females are known to have a significant survival advantage in IPF [7], and this may be an additional indication of the mechanism behind this.

Lung \( T_1 \) correlated negatively with age, but sex was not found to predict \( T_1 \) in this population composed mostly of ILD patients. This can be compared with the findings of Kindval et al. [42] who found a negative correlation between lung \( T_1 \) and age in female healthy volunteers. Lung \( T_1 \) also linearly correlated with FVC\% when age is considered, and IPF patients present higher lung \( T_1 \) than the rest of the population. Stadler at al found that fibrotic patients in inspiration presented higher \( T_1 \) than emphysema patients [43], and this may explain the latter finding since the IPF population tends to present more as a fibrotic phenotype than other ILD types, and the acquisition was done in free breathing, so with inflated lungs.

The relationships found at visit 1 between imaging biomarkers and PFTs were mostly lost when considering short-term (6 weeks or 6 months) longitudinal changes. Among the \( ^{129}\text{Xe} \) ventilation biomarkers, the ventilated volume Xe-VVF change was significantly correlated only with age. Lung \( T_1 \) changes were negatively correlated with the change in FEV_1\%. Since the change in PFTs between visits was small, this may explain the lack of correlation between imaging biomarkers and PFTs.

A limitation of this work is that, as a sub study, it was not adequately powered to address every question discussed here. Furthermore, the subjects were not uniformly distributed across the disease groups, reflecting challenges with the recruitment of rarer ILD variants. It is therefore possible that a larger study would uncover additional relationships. Subacute hypersensitivity pneumonitis is characterised by air trapping [44], so this subgroup may have been expected to have lower ventilated volume fractions than other ILD subtypes, but this was not found in this population. Also, the age differences between healthy volunteers and the IPF group is a potential issue. Potential further development within the field may also lead to a growing understanding.
Fig. 7. Boxplot of pulmonary proton MR biomarkers at visit 1 in the studied population, split in ILD subgroups. (*) $p < 0.05$, (**) $p < 0.01$. $\Delta pO_2$: change in oxygen partial pressure, $\tau_{up}$: wash-in rate of oxygen, OE-VVF: percentage of ventilated volume as calculated by oxygen-enhanced MRI, OE-VHI: ventilation heterogeneity index as calculated by oxygen-enhanced MRI, $T_1$: inversion recovery $T_1$. 

M. Tibiletti et al.
of the imaging techniques and further insights into disease pathology, physiology and progression may develop over time with increasing experience and patient numbers.

Another limitation regards the observational nature of the study. Subjects received varied treatment regimens before and after recruitment, due to the different origin of ILD and a wide range of disease severity. Also, the timing of the baseline study visits in relation to the management of non-IPF subjects varied between the two centres. It is possible that the variation in the patient management in the study could have influenced the longitudinal changes in the biomarkers and a more standardised approach may have resulted in different outcomes. Therefore, the failures of the imaging biomarkers in distinguishing longitudinal changes between the group receiving or not pharmacological treatment must be interpreted cautiously.

Another limitation of our work is its focus on imaging techniques that are centred around ventilation imaging. There exist alternative techniques, such as spectroscopy and diffusion-weighted 129Xe MRI [9,14,45] or 1H perfusion MRI [46], which are capable of probing other lung properties such as gas exchange, lung microstructure and blood perfusion. Such measurements provide a promising alternative class of imaging biomarkers for ILD.

5. Conclusions

In conclusion, none of the global measurements investigated in this study were able to differentiate between ILD subtypes, suggesting that ventilation-only biomarkers are not indicated for this task. Limited but progressive loss of ventilated volume as measured by 129Xe-MRI may be present as disease progresses, but no ventilation biomarker investigated in this study is a good candidate for monitoring longitudinal changes in ILD. 129Xe ventilation biomarkers correlated strongly with age and TLCO at visit 1, but the correlations were mostly lost when considering short-term (6 weeks or 6 months) longitudinal changes. Both OE-MRI and 129Xe MRI revealed more spatially heterogeneous ventilation in DI-ILD and IPF.


