The assessment of short- and long-term changes in lung function in cystic fibrosis using [129]Xe MRI

Laurie J. Smith [1,2], Alex Horsley [1,3], Jody Bray[1], Paul J.C. Hughes [1], Alberto Biancardi[1], Graham Norquay[1], Martin Wildman[4], Noreen West[2], Helen Marshall [1] and Jim M. Wild[1]

Affiliations: [1]POLARIS, Imaging Sciences, Dept of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK. [2]Sheffield Children's Hospital NHS Foundation Trust, Sheffield, UK. [3]Respiratory Research Group, Division of Infection, Immunity and Respiratory Medicine, University of Manchester, Manchester, UK. [4]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK.

Correspondence: Jim M. Wild, POLARIS, Imaging Sciences, Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK. E-mail: j.m.wild@sheffield.ac.uk

ABSTRACT

Introduction: Xenon-129 ([129]Xe) ventilation magnetic resonance imaging (MRI) is sensitive to detect early cystic fibrosis (CF) lung disease and response to treatment. [129]Xe-MRI could play a significant role in clinical trials and patient management. Here we present data on the repeatability of imaging measurements and their sensitivity to longitudinal change.

Methods: 29 children and adults with CF and a range of disease severity were assessed twice, a median (interquartile range (IQR)) of 16.0 (14.4–19.5) months apart. Patients underwent [129]Xe-MRI, lung clearance index (LCI), body plethysmography and spirometry at both visits. 11 patients repeated [129]Xe-MRI in the same session to assess the within-visit repeatability. The ventilation defect percentage (VDP) was the primary metric calculated from [129]Xe-MRI.

Results: At baseline, mean±SD age was 23.0±11.1 years and forced expiratory volume in 1 s (FEV_1) z-score was −2.2±2.0. Median (IQR) VDP was 9.5 (3.4–31.6)% and LCI was 9.0 (7.7–13.7). Within- and inter-visit repeatability of VDP was high. At 16 months there was no single trend of [129]Xe-MRI disease progression. Visible [129]Xe-MRI ventilation changes were common, which reflected changes in VDP. Based on the within-visit repeatability, a significant short-term change in VDP is >±1.6%. For longer-term follow-up, changes in VDP of up to ±7.7% can be expected, or ±4.1% for patients with normal FEV_1. No patient had a significant change in FEV_1; however, 59% had change in VDP >±1.6%. In patients with normal FEV_1, there were significant changes in ventilation and in VDP.

Conclusions: [129]Xe-MRI is a highly effective method for assessing longitudinal lung disease in patients with CF. VDP has great potential as a sensitive clinical outcome measure of lung function and end-point for clinical trials.

Introduction

For people with cystic fibrosis (CF), advances in treatments and patient management have significantly increased expected survival. These advances have greatly improved the lung health of patients, and now the median forced expiratory volume in 1 s (FEV_1) for UK patients aged <18 years is well preserved at 88% predicted [1]. However, it is also well accepted that a value for FEV_1 within the range of normal does not necessarily mean that the patient's lung function is truly normal [2, 3]. With highly effective cystic fibrosis transmembrane conductance regulator (CFTR) modulator therapies being increasingly administered to patients with an FEV_1 within the normal range, more sensitive outcome methods for assessing lung function are required.

Hyperpolarised gas ventilation magnetic resonance imaging (MRI) using either helium-3 (^{3}He) or xenon-129 (^{129}Xe) provides a direct visual and quantitative assessment of the distribution of ventilation within the lung in three dimensions [4]. Ventilation abnormalities are clearly identified as areas of signal deficit and are termed ventilation defects. The ventilation heterogeneity seen on MRI can be quantified using different approaches to assess the degree of abnormality. The metric most widely used is the ventilation defect percentage (VDP), which quantifies the proportion of the image without any ventilation present. A major strength of ventilation MRI is the ability to measure individual ventilation defects, which allows for small regional changes in lung function to be assessed [5].

Previous studies in CF populations using ^{3}He showed that ventilation MRI is highly sensitive to detect early lung disease in subjects with normal values for FEV_1 [6–12], lung clearance index (LCI) [5, 6] and computed tomography imaging [6]. The latter two methods are already recognised as sensitive methods for the detection of early lung disease [3, 13]. A previous study of young patients with mild CF lung disease showed that ^{3}He MRI was sensitive to longitudinal changes in lung function that were largely undetected by FEV_1 and LCI [14]. In more recent years, ventilation MRI research has moved from ^{3}He towards ^{129}Xe due to the lower cost and greater availability of ^{129}Xe. Recent studies utilising ^{129}Xe MRI in CF have found that it is well tolerated [15] and is sensitive to detect early lung disease [16, 17]. In addition, ^{129}Xe MRI has been shown to be sensitive to treatment response to pulmonary exacerbation in children with CF, with ^{129}Xe VDP showing a larger treatment response than both LCI and FEV_1 [18].

Hyperpolarised gas ventilation MRI is therefore an attractive method for assessing CF lung disease, but more data are required on the background longitudinal changes seen in stable CF lung disease. This includes a systematic assessment of the intrinsic technical repeatability of the measurement in this population, which preliminary data suggest is promising [19], as well as the pathophysiological variability seen in stable disease. Therefore, in this study we aimed to assess the potential of ^{129}Xe MRI as a quantitative outcome measure of lung health and a possible candidate end-point for clinical trials and patient management. In order to do this, we assessed the short- and long-term repeatability of imaging (VDP) in a cohort of children and adults with CF and a range of lung disease. In addition, we aimed to better understand the longitudinal changes in lung function on ^{129}Xe MRI in comparison to LCI and FEV_1 in those patients with an FEV_1 within the normal range.

Methods

This prospective study recruited adults and children (aged >5 years) with CF from three specialist CF centres in the UK (Sheffield Children's Hospital and Northern General Hospital, Sheffield, UK and Manchester Adult CF Centre, Manchester, UK). For inclusion, patients had to be clinically stable for 4 weeks prior to assessment (defined as free from intravenous antibiotics or any hospital stay within 4 weeks) and have an FEV_1 >30% pred within the previous 6 months. This study was approved by the Yorkshire and Humber – Leeds West research ethics committee (16/YH/0339). Parents/guardians of children and all adult patients provided written informed consent.

Lung imaging

Hyperpolarised ^{129}Xe ventilation MRI of the lungs was performed at a single site (Sheffield) on a 1.5 T GE HDx scanner (GE, Milwaukee, WI, USA) using previously described protocols [20]. ^{129}Xe was polarised using a bespoke spin exchange optical pumping polariser [21], under a UK Medicines and Healthcare Products Regulatory Agency manufacturing specials regulatory licence (MS-18739). Ventilation imaging was acquired at a lung volume of end-inspiratory tidal volume, by inhaling a volume of ^{129}Xe titrated with a balance of medical-grade nitrogen, from a lung volume of functional residual capacity. The total inhaled gas volume ranged from 0.4 to 1.0 L and was calculated based on the subject's height. For ^{129}Xe analysis a hydrogen-1 (^{1}H) anatomical image was performed in a separate breath-hold (immediately prior to the ^{129}Xe image) in order to calculate the thoracic cavity volume. For quantitative analysis, the ^{1}H and ventilation images were segmented using a semi-automated method [22], from which the VDP and the ventilation heterogeneity index (VH_I), which reflects the heterogeneity of ventilated voxels within

ventilated lung regions, were calculated as described previously [5]. Further details of image acquisition and processing methods can be found in the supplementary material.

In order to assess the within-visit technical repeatability of ^{129}Xe-MRI, ^{129}Xe imaging was repeated in a subgroup of patients within 15 min of the initial baseline measurement and without the subject leaving the scanner. The same imaging protocol, respiratory manoeuvres and volume of gas were used for both scans.

Lung physiology

Patients performed multiple-breath washout on the same day and at the same centre as imaging using an open-circuit Innocor gas analyser (Pulmotrace, Glamsberg, Denmark) using 0.2% sulfur hexafluoride [23]. LCI was calculated from the average of three trials as recommended [24]. Body plethysmography was performed using a PFT Pro (Vyaire, Basingoke, UK) according to guidelines [25], in order to calculate the ratio of residual volume (RV) to total lung capacity (TLC). Finally, spirometry was performed according to guidelines [26] and expressed as z-scores [27]. Either MRI or LCI was performed first, followed by the other. Spirometry was always performed last.

The assessments of ^{129}Xe ventilation MRI, multiple-breath washout and spirometry were then repeated at a second stable visit using the methods described.

Statistical analysis

Data were analysed using Prism version 8.0 (GraphPad, San Diego, CA, USA) and SPSS statistics version 26.0 (IBM, Armonk, NY, USA). Normal distribution was assessed using the Shapiro–Wilk test. Data are expressed as mean±SD for normally distributed data, and median (interquartile range (IQR)) for nonparametric data. Within-visit repeatability of ^{129}Xe-MRI was assessed from the 95% limits of agreement (LoA) of a Bland–Altman analysis of the repeat measurements. Wilcoxon signed-rank test and Bland–Altman analysis were used to compare MRI and lung function metrics between the baseline and follow-up study visits. Within- and inter-visit repeatability was calculated using the intraclass correlation coefficient (ICC). Spearman correlation analysis was used to compare the change in different metrics. Sample size power calculations were calculated for different effect sizes based on the longitudinal data [28]. Statistical significance was set at $p < 0.05$.

In order to assess whether a clinically significant change in VDP had occurred over time, firstly the Bland–Altman LoA from the within-visit repeatability measurements was used as a minimal threshold. Secondly, a further threshold to represent a clinically significant change in absolute VDP was set at ±3%; this threshold represents the mean absolute change in ^{129}Xe VDP in response to treatment of an exacerbation of CF lung disease [18]. In order to compare the significance of change in VDP, similar thresholds for short- and long-term repeatability were applied for LCI and FEV_1. Repeatability of ±10% in LCI has been shown for healthy volunteers [23], while longitudinal changes of ±20% have been reported in clinically stable CF patients [29]. For FEV_1, a within-patient longitudinal change of >±10% is deemed significant [30], while the short-term change of FEV_1 in patients with CF is approximately ±5% [31, 32].

Results

29 children and adults with CF were assessed on two occasions, a median of 16 months apart. At baseline, patients were aged between 6 and 47 years. Baseline demographics, MRI metrics and lung function are detailed in table 1. All but one patient had visible ventilation defects present at both study visits. 14 (48%) patients had a normal FEV_1 value (>−1.64 z-score) at both baseline and follow-up. Three (10%) patients at baseline, and seven (24%) at follow-up had normal LCI values.

Within-visit repeatability of ^{129}Xe MRI

11 (35%) patients performed repeat ^{129}Xe-MRI within 15 min of the baseline scan. Median (IQR) age 23.7 (17.7–33.2) years, baseline VDP 7.3 (2.5–30.8)%, LCI 8.3 (7.3–14.0), FEV_1 −2.4 (−2.8−−0.5) z-score. There was no significant difference in VDP between scans and good repeatability with a bias of 0.2% and 95% LoA −1.4–1.8%. This represents the intrinsic technical repeatability of the measurement *in vivo* assuming no true change in underlying lung ventilation. Based on this analysis a threshold of absolute change in VDP of ±1.6% was used in part to assess ^{129}Xe VDP longitudinal change. (Repeatability for VH_I can be found in figure 5, alongside the Bland–Altman plots for VDP.) For VH_I, there was again minimal bias (−0.6), with 95% LoA −2.5–1.2%. The within-visit ICC for VDP was excellent at 0.99 (95% CI 0.99–1.0), and was 0.96 (95% CI 0.84–0.99) for VH_I.

Longitudinal change in ^{129}Xe MRI

All 29 patients successfully repeated ^{129}Xe MRI and lung function testing at a second visit, after a median (IQR) interval of 16.0 (14.4–19.5) months. There was no single pattern of disease progression in the

TABLE 1 Patient demographics

	Baseline visit	Follow-up visit	Absolute change[#] Δ	Relative change[¶] %	ICC[+] (95% CI)	Bland–Altman bias[#] (LoA)
Demographics						
Patients n (% female)	29 (52)					
Age years	23.0±11.1	24.3±11.1				
Height cm	160.3±16.2	162.9±14.3				
Weight kg	54.7±17.4	56.6±17.0				
^{129}Xe MRI						
VDP %	9.5 (3.4–31.6)	10.5 (3.0–29.4)	0.5 (−1.8–2.4)	8.2 (−13.9–35.5)	0.97 (0.94–0.99)	0.8 (−7.0–8.5)
VH$_I$ %	14.2 (10.3–17.7)	12.6 (9.7–18.3)	−0.1 (−0.9–1.1)	−1.1 (−7.1–11.0)	0.89 (0.80–0.95)	0.3 (−3.7–4.3)
Pulmonary function						
FEV$_1$ z-score	−2.2±2.0	−2.3±2.0	−0.0±0.4	0.2 (−14.5–12.0)	0.98 (0.96–0.99)	−0.1 (−0.8–0.7)
FEV$_1$ % predicted	72.1±25.6	71.3±25.5	−0.8±5.0	−1.1±7.9	0.98 (0.96–0.99)	−0.8 (−10.5–8.9)
LCI	9.0 (7.7–13.7)	8.9 (7.5–14.5)	0.3 (−0.9–1.1)	4.3 (−7.9–10.2)	0.95 (0.90–0.98)	0.3 (−2.2–2.8)
RV/TLC %	35.3 (26.0–47.1)	34.4 (25.3–48.7)	0.0 (−2.5–2.7)	0.0 (−6.4–7.9)	0.96 (0.91–0.98)	0.1 (−7.3–7.6)

Data are presented as mean±SD or median (interquartile range), unless otherwise stated. ICC: intraclass correlation coefficient; LoA: limits of agreement; ^{129}Xe: xenon-129; MRI: magnetic resonance imaging; VDP: ventilation defect percentage; VH$_I$: ventilation heterogeneity index; FEV$_1$: forced expiratory volume in 1 s; LCI: lung clearance index; RV: residual volume; TLC: total lung capacity. [#]: difference between follow-up and baseline; [¶]: percentage change from baseline; [+]: ICC between follow-up and baseline.

cohort and no lung function or MRI metric demonstrated a statistically significant group change between visits. Instead, significant inter-subject variation was seen in the degree and direction of change in ventilation distribution on ^{129}Xe MRI. For many patients there were clear and often large visible changes in the distribution of ventilation, independent of underlying disease severity (figures 1–3). Figure 1 shows eight example images from patients, all of whom had FEV$_1$ in the normal range, where there was a change in the distribution of ventilation and in VDP, but without significant change in FEV$_1$ or LCI (see also figures 2 and 3).

Longitudinal change relative to baseline for VDP, LCI and FEV$_1$ are shown in figure 4. Overall, 17 (59%) patients had an increase (worsening) in VDP at follow-up, which correlated with the visual image analysis (figures 4 and 5). 17 (59%) patients had a change in ^{129}Xe VDP of >±1.6%, while nine (31%) also had an absolute change in VDP >±3%. In comparison, 13 (45%) had a relative change in LCI >10% from baseline, but only two (7%) had a relative change >±20%. For FEV$_1$, 10 (34%) had an absolute change in FEV$_1$ of >±5% pred and no patients had an absolute change >±10%. Of the nine patients with a change in VDP >±3%, no patient had a corresponding significant change in LCI or FEV$_1$. Of all the metrics, ^{129}Xe VDP had the highest median relative change over time (8.2%).

Inter-visit repeatability
The inter-visit ICC for ^{129}Xe VDP was excellent at 0.97, which was similar to FEV$_1$ (0.98) and higher than LCI (0.95), RV/TLC (0.96) and ^{129}Xe VH$_I$ (0.89) (table 1). The change in VDP tended to be larger for those with higher baseline VDP. When only patients with normal FEV$_1$ were considered (and therefore with lower values for VDP) the 95% LoA fell from −6.9% to 8.5% for the whole cohort to −4.3% to 4.0% (n=14) (figure 5).

Correlation of the changes in metrics over time
The absolute or relative change in VDP was not correlated with absolute or relative change in FEV$_1$, LCI or RV/TLC. In contrast, the absolute and the relative changes in VH$_I$ and LCI were significantly correlated with each other (r=0.68, p<0.001 and r=0.73, p<0.001, respectively). There was no correlation in the change in either FEV$_1$ or RV/TLC with the other metrics. In addition, there was no relationship between the magnitude of change in VDP with age or underlying lung disease as measured at baseline.

Sample size power calculations
With a view to using VDP from ^{129}Xe MRI as an intervention outcome marker, sample size calculations for four different effect sizes and three populations were derived. Effect sizes include the minimal change of 1.6% in VDP, the mean change of 3% seen with *i.v.* antibiotics and 5% and 10% change. Population mean±SD are taken from the whole-cohort data (representing a mixed CF population with a wide range of
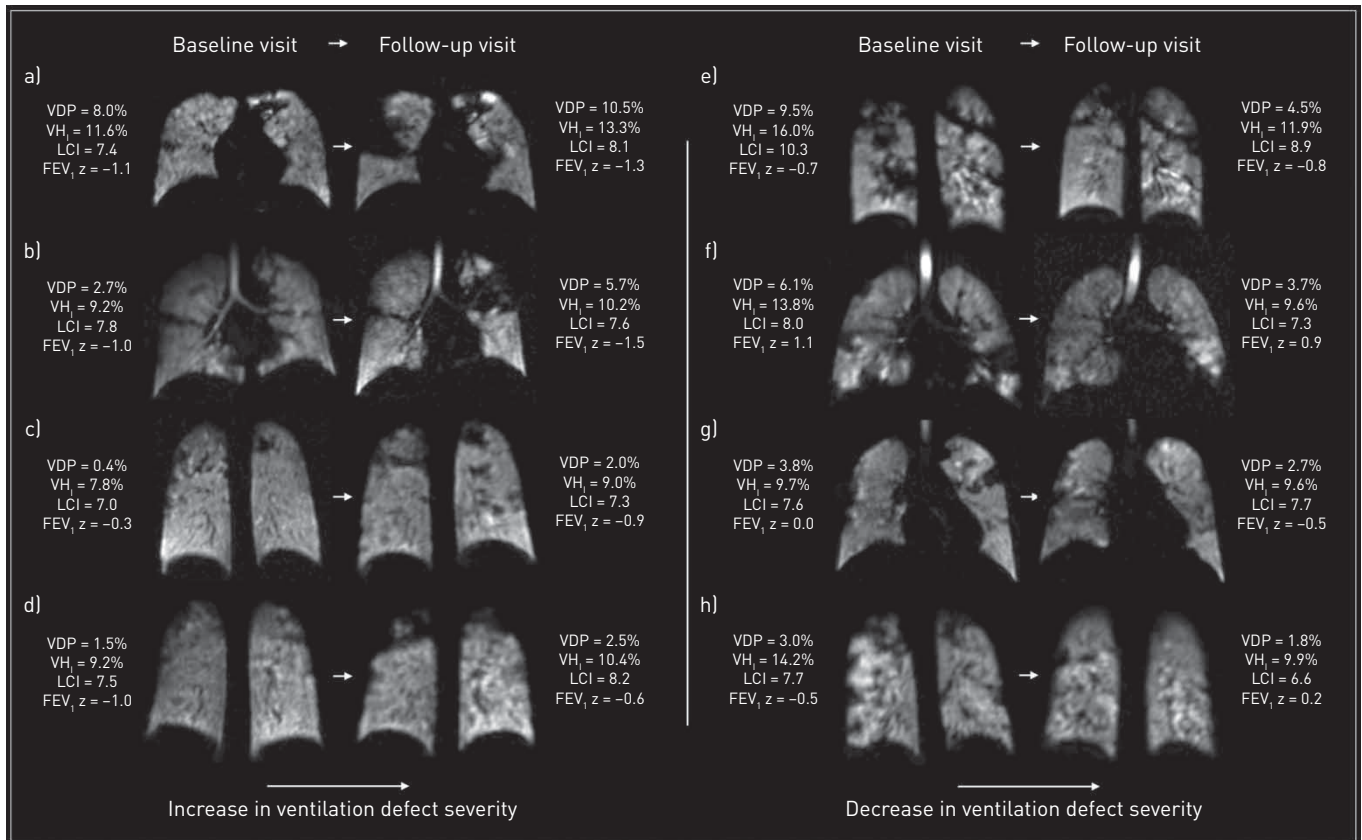
FIGURE 1 Eight longitudinal xenon-129 magnetic resonance imaging examples from patients with normal forced expiratory volume in 1 s (FEV_1) values. a–d) Patients with both a visible increase in ventilation defect severity and increase in ventilation defect percentage (VDP) over time; e–h) patients with both a visible decrease in ventilation defect severity and in VDP over time. VH_I: ventilation heterogeneity index; LCI: lung clearance index.
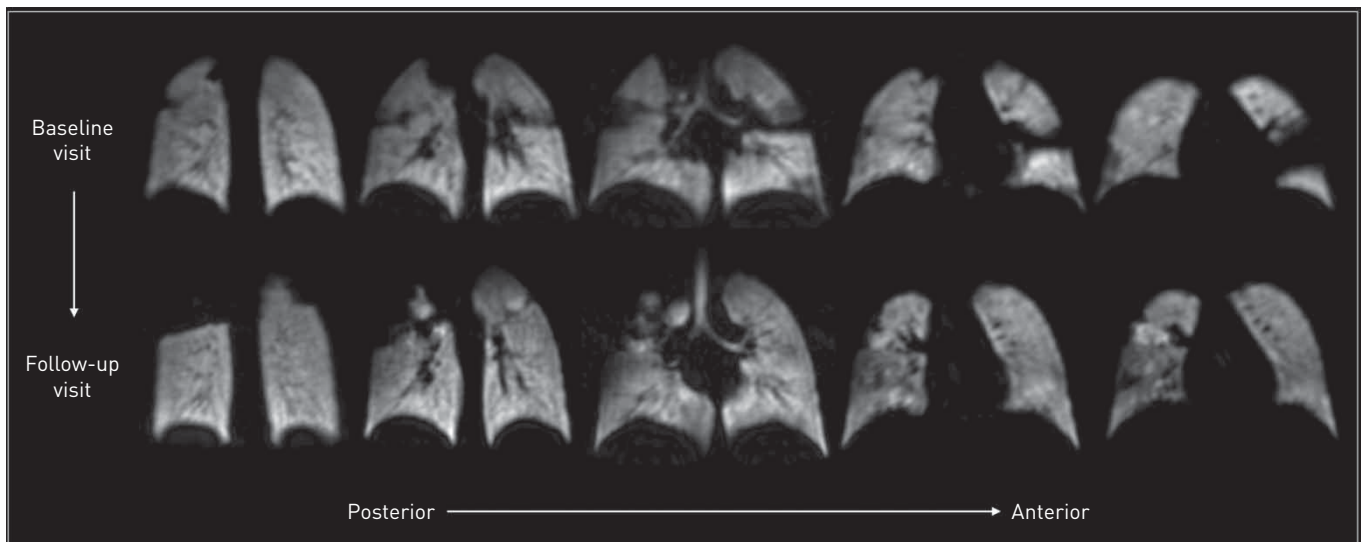

FIGURE 2 Longitudinal xenon-129 magnetic resonance imaging in a patient with normal forced expiratory volume in 1 s (FEV_1) and lung clearance index (LCI). At follow-up there was a change in VDP from 5.4% to 6.7%. At follow-up, this patient's FEV_1 and LCI were similar to baseline, FEV_1 z-score changed from −0.7 to −1.1 and LCI changed from 6.7 to 7.0. The longitudinal ventilation change in this patient differs between the two lungs. The ventilation defects in the left lung have disappeared and lung disease appears to have improved over time. In contrast, the right lung shows a significant increase/worsening in the degree of ventilation abnormalities present over time. This example highlights the importance of utilising both the qualitative and quantitative image information together and the potential for analysing regional lung disease.
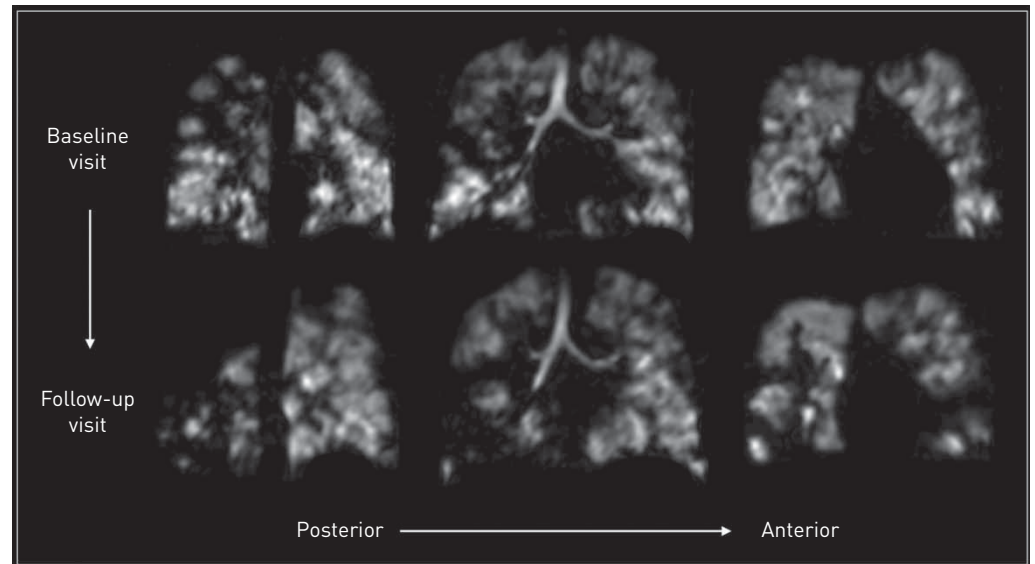
FIGURE 3 Longitudinal xenon-129 magnetic resonance imaging in a patient with abnormal forced expiratory volume in 1 s (FEV_1). This patient had a large and significant worsening in ventilation abnormalities over time, reflected by an increase in ventilation defect percentage (VDP) from 16.6% to 27.3%. Despite these ventilation changes, the FEV_1 z-score change was from −4.0 to −4.4 and the LCI change was from 12.9 to 13.6. In more advanced lung disease, as with this example, ventilation imaging can provide a detailed assessment of longitudinal change in lung function that conventional metrics are not sensitive to detect.
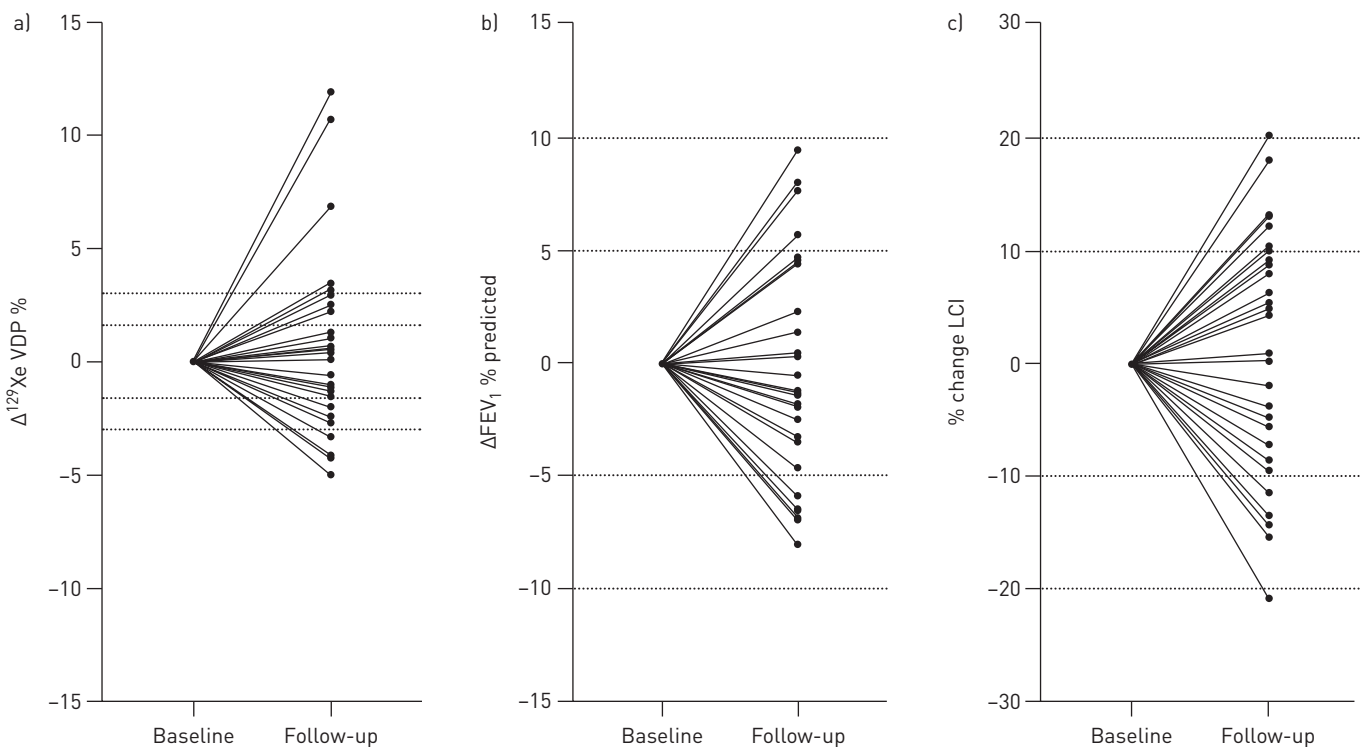

FIGURE 4 The change over time from baseline for a) ventilation defect percentage (VDP), b) forced expiratory volume in 1 s (FEV_1) and c) lung clearance index (LCI). The data are displayed so that the baseline data point always reads zero. For both VDP and FEV_1 the absolute change is shown, while for LCI the relative change is shown. For VDP the inner dotted lines at ±1.6% represent the limits of agreement of the same-day repeatability, while the outer dotted lines at ±3% represent the mean treatment response seen in the study by RAYMENT et al. [18]. For FEV_1 the inner dotted lines at ±5% represent a low threshold of repeatability as a means of comparison with the within-visit repeatability of xenon-129 (^{129}Xe) magnetic resonance imaging. The outer dotted lines at ±10% represent the mean treatment response seen in the study by HORSLEY et al. [30]. For LCI, the inner dotted lines at ±10% represent a low threshold again in order to compare to the within-visit repeatability of ^{129}Xe VDP. The outer dotted lines at ±20% represent a clinical threshold for a significant change. Out of the population of 29 patients, nine had a change in VDP greater than the wider threshold, as opposed to none for FEV_1 and two for LCI.
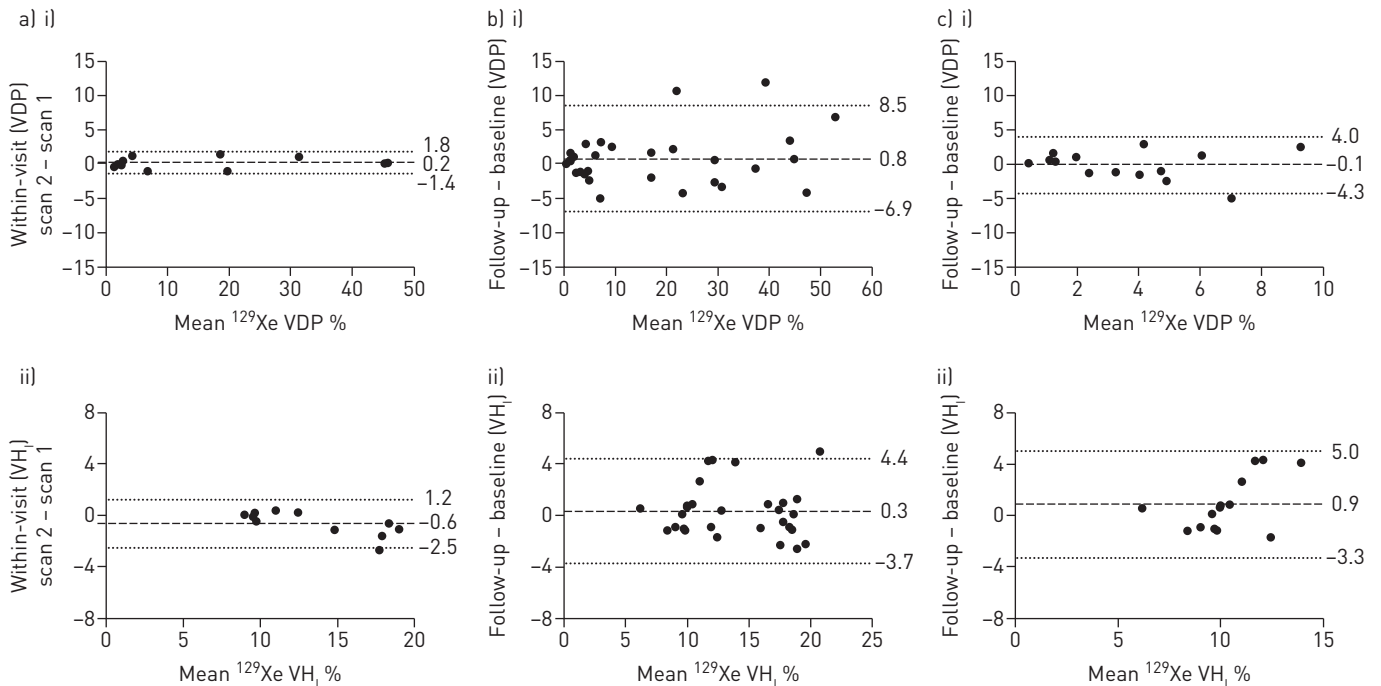
FIGURE 5 Bland–Altman plots comparing ventilation defect percentage (VDP) and ventilation heterogeneity index (VH$_I$) at the same visit and at follow-up when compared to baseline. a) Within-visit patient group for i) VDP and ii) VH$_I$; b) whole cohort when followed-up over time for i) VDP and ii) VH$_I$; c) patients with normal forced expiratory volume in 1 s (FEV$_1$) when followed-up over time for i) VDP and ii) VH$_I$. The dashed lines on both plots indicate the bias and the dotted lines indicate the 95% limits of agreement with the values detailed. When considering VDP, the limits of agreement are considerably narrower for the within-visit group when compared to the follow-up over time. Considering only the long-term follow-up, the limits of agreement are considerably narrower when only the patients with normal FEV$_1$ are considered, when compared to the whole cohort.

disease severity) and, separately, only those with normal-range FEV$_1$. The results are presented in table 2. This emphasises the importance of the baseline variability and appropriate population selection, but also shows the low numbers that are potentially required to detect significant change. For example, a 3% change could be detected with a power of 90% in a study population of 11 patients with CF and normal FEV$_1$.

Discussion

The data reported in this study are the first longitudinal assessment of patients with CF using ^{129}Xe lung ventilation MRI. In this study we demonstrate that 1) ^{129}Xe MRI VDP has high within-visit repeatability; 2) a qualitative and quantitative approach to image analysis is complementary in assessing CF lung disease; and 3) in patients with a preserved FEV$_1$, as well as those with more advanced disease, VDP demonstrates changes in ventilation distribution in patients where FEV$_1$ and LCI do not show significant change.

TABLE 2 Sample size power calculations for different effect sizes, based on xenon-129 (^{129}Xe) ventilation defect percentage (VDP) and different patient populations

| | Subjects n | Standard deviation of ΔVDP# | Effect size | | | | | | | |
| | | | 1.6% | | 3% | | 5% | | 10% | |
			80%	90%	80%	90%	80%	90%	80%	90%
Subgroup, within-visit repeatability¶	11	0.82	5	6	2	2	1	1	1	1
Whole CF cohort, longitudinal repeatability+	29	3.93	95	127	27	37	10	13	3	4
CF cohort with normal FEV$_1$, longitudinal repeatability§	14	2.11	28	37	8	11	3	4	1	1

CF: cystic fibrosis; FEV$_1$: forced expiratory volume in 1 s. #: the standard deviation of the difference in VDP between two time-points; ¶: patients who underwent same-day ^{129}Xe repeatability scans; +: the complete cohort with two ^{129}Xe images ~16 months apart; §: the same cohort, but with only those included who had normal FEV$_1$.

There is a growing body of evidence that ventilation MRI provides valuable and detailed insights into the underlying lung function of patients with CF that is not detected by other methods [5–12, 14, 16–18, 33]. Previous studies have shown that ventilation MRI is highly sensitive to early lung disease and, in the case of ^{3}He, is repeatable [34, 35] and sensitive to disease progression [14]. The data presented in this study add to this evidence base by demonstrating that ^{129}Xe MRI is also highly sensitive to detect longitudinal changes in CF lung disease. Unlike the previous study from our group [14], performed in children with mild CF lung disease using ^{3}He MRI, here we did not see one single pattern of disease progression. Of the cohort reported here, 59% of patients had evidence of increased (worsening) VDP, while the remaining patients showed improvements in ventilation. This is not surprising given that this is a broad cohort of patients, in terms of age and disease severity, although there was no relationship between the magnitude of change in VDP with either age or disease severity. Ventilation defects caused by mucus obstruction will not necessarily remain stable. Thus, some visible defects will represent short-term reversible obstructions while others may be caused by underlying disease progression and airway narrowing. Figure 1 shows how in patients with normal FEV$_1$, ^{129}Xe MRI is able to detect early disease-related changes, as we previously reported with ^{3}He MRI. Furthermore, figure 1 shows that changes in lung ventilation can be measured on ^{129}Xe MRI that are not necessarily detected by LCI. This is a particularly important finding in the era of new and expensive CFTR-modulating therapies, where there is a need to be able to measure clinical response to therapies even in those with apparently normal lung function. ^{129}Xe MRI VDP is a metric that may provide this detail, as has been shown previously for ^{3}He MRI in the assessment of ivacaftor [33]. In addition, there is also increasing evidence for the application of ^{1}H structural MRI in the clinical assessment of CF lung disease [36–39]. ^{1}H MRI can be performed at the same visit as ^{129}Xe MRI, allowing for the combined assessment of lung structure and function when measured together.

Inter-visit repeatability is affected by sources of both intrinsic (technical) and physiological variability, as well as true disease progression. For VDP, a change of >±1.6% is greater than the inter-subject intrinsic repeatability of the measurement, and potentially represents a lower threshold for significant change. For comparison, a median change of >±3% in VDP in response to *i.v.* antibiotics should represent a clinically significant degree of change [18]. The true threshold for clinically relevant change in VDP therefore probably lies between these limits, but cannot be determined more precisely from the data available. However, over longer time courses, CF patients have natural fluctuations in mucus plugging and symptoms, which are separate from underlying disease progression. We have shown that over 16 months, a change in VDP of up to ±7.7% is seen in CF patients considered to be clinically stable and without obvious disease progression by other lung function metrics; the change is less (at ±4.1%) in those with a preserved FEV$_1$.

The findings that lung function metrics on average are unchanged with time are consistent with longitudinal analyses of patients with CF using LCI, where minimal longitudinal change was reported [40–43]. Our findings, in addition, highlight that patients with CF are as likely to improve clinically as they are to have deteriorating ventilation heterogeneity during observational follow-up. Despite this, figures 1 and 3–5 show that patients often had subclinical changes in VDP without significant change in FEV$_1$ or LCI. It is likely that some of these changes we have seen in ventilation are transient and some are the precursor to exacerbation and potentially irreversible ventilation changes.

In this study we highlight how VDP can sensitively track changes in underlying lung function in patients with preserved FEV$_1$, which correspond to visual changes on the ventilation images (figures 1 and 2). In order to assess this specific patient population in clinical trials, relatively large sample sizes are required to measure modest treatment effects when FEV$_1$ is the primary outcome [44]. LCI has been used as an alternative outcome in more recent studies [45], which allows for smaller sample sizes in patients with normal FEV$_1$. However, ^{129}Xe VDP has high repeatability and low standard deviation and large effect sizes can be measured, which is reflected in the relatively small sample sizes required to measure the different reported effect sizes.

^{129}Xe VDP is an attractive potential outcome measure/end-point in both clinical trials and clinical management. A strength of ^{129}Xe MRI is not only that can we produce summary whole-lungs metrics like VDP that are more sensitive than LCI and FEV$_1$ at detecting early lung disease, but the images themselves also contain more detailed regional functional information [5]. Therefore, it is possible to detect clinically relevant regional change even in the face of apparently unchanged lung physiology tests [5, 46]. This applies both to detecting disease progression in clinical practice over time courses like the one described in this study, as well as detecting much shorter term improvements due to therapeutic interventions [18, 33]. In order to generate quantitative regional metrics of lung physiology from ventilation MRI, future work should focus on reliable parameterisation of regional ventilation heterogeneity to further improve the clinical utility of ventilation MRI.

We recognise that hyperpolarised gas ventilation MRI is not currently available to many CF centres. Estimates of regional lung function may be acquired indirectly, without the use of inhaled contrast agents

using time-resolved ^{1}H MRI techniques [47, 48] or by using contrast enhanced perfusion MRI [38]. These techniques are promising and may provide an alternative, more widely accessible method to the wider CF community. A further limitation of this study is the lack of detailed clinical data to cover the period between visits, which may have helped explain some of the changes seen. We also acknowledge that this is a single-centre analysis, which may have an impact on the data; however, a recent study reported the high repeatability of ^{129}Xe VDP in a multicentre setting [49], which highlights the potential of VDP as an end-point in multicentre studies.

In conclusion, ^{129}Xe MRI is a highly effective method for the assessment of CF lung disease. In this study, ^{129}Xe VDP has high within-visit repeatability and measures underlying changes in lung function that are not necessarily detected by other methods. Measuring small changes in lung function in a patient population with increasingly normal, preserved spirometry values is challenging but highly relevant. ^{129}Xe ventilation MRI can both qualitatively and quantitatively meet this requirement and should therefore be considered as a future end-point for clinical trials and patient management.

References

1   Cystic Fibrosis Trust. UK Cystic Fibrosis Registry, Annual Data Report 2018. 2019. Available from: www.cysticfibrosis.org.uk
2   Brody AS, Tiddens HA, Castile RG, et al. Computed tomography in the evaluation of cystic fibrosis lung disease. Am J Respir Crit Care Med 2005; 172: 1246–1252.
3   Aurora P, Gustafsson P, Bush A, et al. Multiple breath inert gas washout as a measure of ventilation distribution in children with cystic fibrosis. Thorax 2004; 59: 1068–1073.
4   Woods JC, Wild JM, Wielputz MO, et al. Current state of the art MRI for the longitudinal assessment of cystic fibrosis. J Magn Reson Imaging 2020; 52: 1306–1320.
5   Smith LJ, Collier GJ, Marshall H, et al. Patterns of regional lung physiology in cystic fibrosis using ventilation magnetic resonance imaging and multiple-breath washout. Eur Respir J 2018; 52: 1800821.
6   Marshall H, Horsley A, Taylor CJ, et al. Detection of early subclinical lung disease in children with cystic fibrosis by lung ventilation imaging with hyperpolarised gas MRI. Thorax 2017; 72: 760–762.
7   Bannier E, Cieslar K, Mosbah K, et al. Hyperpolarized 3He MR for sensitive imaging of ventilation function and treatment efficiency in young cystic fibrosis patients with normal lung function. Radiology 2010; 255: 225–232.
8   Woodhouse N, Wild JM, van Beek EJ, et al. Assessment of hyperpolarized 3He lung MRI for regional evaluation of interventional therapy: a pilot study in pediatric cystic fibrosis. J Magn Reson Imaging 2009; 30: 981–988.
9   van Beek EJ, Hill C, Woodhouse N, et al. Assessment of lung disease in children with cystic fibrosis using hyperpolarized 3-helium MRI: comparison with Shwachman score, Chrispin–Norman score and spirometry. Eur Radiol 2007; 17: 1018–1024.
10   McMahon CJ, Dodd JD, Hill C, et al. Hyperpolarized 3helium magnetic resonance ventilation imaging of the lung in cystic fibrosis: comparison with high resolution CT and spirometry. Eur Radiol 2006; 16: 2483–2490.
11   Mentore K, Froh DK, de Lange EE, et al. Hyperpolarized HHe 3 MRI of the lung in cystic fibrosis: assessment at baseline and after bronchodilator and airway clearance treatment. Acad Radiol 2005; 12: 1423–1429.
12   Koumellis P, van Beek EJ, Woodhouse N, et al. Quantitative analysis of regional airways obstruction using dynamic hyperpolarized ^{3}He MRI – preliminary results in children with cystic fibrosis. J Magn Reson Imaging 2005; 22: 420–426.
13   de Jong PA, Nakano Y, Lequin MH, et al. Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis. Eur Respir J 2004; 23: 93–97.
14   Smith L, Marshall H, Aldag I, et al. Longitudinal assessment of children with mild cystic fibrosis using hyperpolarized gas lung magnetic resonance imaging and lung clearance index. Am J Respir Crit Care Med 2018; 197: 397–400.

15   Walkup LL, Thomen RP, Akinyi TG, *et al.* Feasibility, tolerability and safety of pediatric hyperpolarized ^{129}Xe magnetic resonance imaging in healthy volunteers and children with cystic fibrosis. *Pediatr Radiol* 2016; 46: 1651–1662.

16   Thomen RP, Walkup LL, Roach DJ, *et al.* Hyperpolarized ^{129}Xe for investigation of mild cystic fibrosis lung disease in pediatric patients. *J Cyst Fibros* 2017; 16: 275–282.

17   Kanhere N, Couch MJ, Kowalik K, *et al.* Correlation of lung clearance index with hyperpolarized ^{129}Xe magnetic resonance imaging in pediatric subjects with cystic fibrosis. *Am J Respir Crit Care Med* 2017; 196: 1073–1075.

18   Rayment JH, Couch MJ, McDonald N, *et al.* Hyperpolarised ^{129}Xe magnetic resonance imaging to monitor treatment response in children with cystic fibrosis. *Eur Respir J* 2019; 53: 1802188.

19   Santyr G, Kanhere N, Morgado F, *et al.* Hyperpolarized gas magnetic resonance imaging of pediatric cystic fibrosis lung disease. *Acad Radiol* 2019; 26: 344–354.

20   Stewart NJ, Chan HF, Hughes PJC, *et al.* Comparison of ^{3}He and ^{129}Xe MRI for evaluation of lung microstructure and ventilation at 1.5 T. *J Magn Reson Imaging* 2018; 48: 632–642.

21   Norquay G, Collier GJ, Rao M, *et al.* ^{129}Xe-Rb spin-exchange optical pumping with high photon efficiency. *Phys Rev Lett* 2018; 121: 153201.

22   Hughes PJC, Horn FC, Collier GJ, *et al.* Spatial fuzzy c-means thresholding for semiautomated calculation of percentage lung ventilated volume from hyperpolarized gas and ^{1}H MRI. *J Magn Reson Imaging* 2018; 47: 640–646.

23   Horsley AR, Gustafsson PM, Macleod KA, *et al.* Lung clearance index is a sensitive, repeatable and practical measure of airways disease in adults with cystic fibrosis. *Thorax* 2008; 63: 135–140.

24   Robinson PD, Latzin P, Verbanck S, *et al.* Consensus statement for inert gas washout measurement using multiple- and single-breath tests. *Eur Respir J* 2013; 41: 507–522.

25   Wanger J, Clausen JL, Coates A, *et al.* Standardisation of the measurement of lung volumes. *Eur Respir J* 2005; 26: 511–522.

26   Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.

27   Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3–95-yr age range: the Global Lung Function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.

28   Noordzij M, Tripepi G, Dekker FW, *et al.* Sample size calculations: basic principles and common pitfalls. *Nephrol Dial Transplant* 2010; 25: 1388–1393.

29   Horsley A, Alrumuh A, Bayfield K, *et al.* The potential of closed circuit lung clearance index (LCI) to provide longitudinal clinical utility in cystic fibrosis (CF). *Eur Respir J* 2019; 54: Suppl. 63, PA332.

30   Horsley AR, Davies JC, Gray RD, *et al.* Changes in physiological, functional and structural markers of cystic fibrosis lung disease with treatment of a pulmonary exacerbation. *Thorax* 2013; 68: 532–539.

31   Stanbrook MB, Corey M, Tullis DE. The repeatability of forced expiratory volume measurements in adults with cystic fibrosis. *Chest* 2004; 125: 150–155.

32   Taylor-Robinson D, Whitehead M, Diderichsen F, *et al.* Understanding the natural progression in %FEV$_1$ decline in patients with cystic fibrosis: a longitudinal study. *Thorax* 2012; 67: 860–866.

33   Altes TA, Johnson M, Fidler M, *et al.* Use of hyperpolarized helium-3 MRI to assess response to ivacaftor treatment in patients with cystic fibrosis. *J Cyst Fibros* 2017; 16: 267–274.

34   O'Sullivan B, Couch M, Roche JP, *et al.* Assessment of repeatability of hyperpolarized gas MR ventilation functional imaging in cystic fibrosis. *Acad Radiol* 2014; 21: 1524–1529.

35   Kirby M, Svenningsen S, Ahmed H, *et al.* Quantitative evaluation of hyperpolarized helium-3 magnetic resonance imaging of lung function variability in cystic fibrosis. *Acad Radiol* 2011; 18: 1006–1013.

36   Stahl M, Wielpütz MO, Ricklefs I, *et al.* Preventive Inhalation of Hypertonic Saline in Infants with Cystic Fibrosis (PRESIS): a randomized, double-blind, controlled study. *Am J Respir Crit Care Med* 2019; 199: 1238–1248.

37   Wielpütz MO, Eichinger M, Wege S, *et al.* Midterm reproducibility of chest magnetic resonance imaging in adults with clinically stable cystic fibrosis and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2019; 200: 103–107.

38   Wielpütz MO, Puderbach M, Kopp-Schneider A, *et al.* Magnetic resonance imaging detects changes in structure and perfusion, and response to therapy in early cystic fibrosis lung disease. *Am J Respir Crit Care Med* 2014; 189: 956–965.

39   Wielpütz MO, von Stackelberg O, Stahl M, *et al.* Multicentre standardisation of chest MRI as radiation-free outcome measure of lung disease in young children with cystic fibrosis. *J Cyst Fibros* 2018; 17: 518–527.

40   Green K, Kongstad T, Skov M, *et al.* Variability of monthly nitrogen multiple-breath washout during one year in children with cystic fibrosis. *J Cyst Fibros* 2018; 17: 242–248.

41   Svedberg M, Gustafsson PM, Robinson PD, *et al.* Variability of lung clearance index in clinically stable cystic fibrosis lung disease in school age children. *J Cyst Fibros* 2018; 17: 236–241.

42   Engberink EO, Ratjen F, Davis SD, *et al.* Inter-test reproducibility of the lung clearance index measured by multiple breath washout. *Eur Respir J* 2017; 50: 1700433.

43   O'Neill K, Tunney MM, Johnston E, *et al.* Lung clearance index in adults and children with cystic fibrosis. *Chest* 2016; 150: 1323–1332.

44   Stanojevic S, Ratjen F. Physiologic endpoints for clinical studies for cystic fibrosis. *J Cyst Fibros* 2016; 15: 416–423.

45   Davies J, Sheridan H, Bell N, *et al.* Assessment of clinical response to ivacaftor with lung clearance index in cystic fibrosis patients with a G551D-CFTR mutation and preserved spirometry: a randomised controlled trial. *Lancet Respir Med* 2013; 1: 630–638.

46   Horn FC, Marshall H, Collier GJ, *et al.* Regional ventilation changes in the lung: treatment response mapping by using hyperpolarized gas MR imaging as a quantitative biomarker. *Radiology* 2017; 284: 854–861.

47   Nyilas S, Bauman G, Sommer G, *et al.* Novel magnetic resonance technique for functional imaging of cystic fibrosis lung disease. *Eur Respir J* 2017; 50: 1701464.

48   Voskrebenzev A, Gutberlet M, Klimeš F, *et al.* Feasibility of quantitative regional ventilation and perfusion mapping with phase-resolved functional lung (PREFUL) MRI in healthy volunteers and COPD, CTEPH, and CF patients. *Magn Reson Med* 2018; 79: 2306–2314.

49   Couch MJ, Thomen R, Kanhere N, *et al.* A two-center analysis of hyperpolarized ^{129}Xe lung MRI in stable pediatric cystic fibrosis: potential as a biomarker for multi-site trials. *J Cyst Fibros* 2019; 18: 728–733.