RESEARCH ARTICLE

Same-Day Repeatability and 28-Day Reproducibility of Xenon MRI Ventilation in Children With Cystic Fibrosis in a Multi-Site Trial

Laura L. Walkup, PhD,[1,2,3*] David J. Roach, PhD,[1] Joseph W. Plummer, PhD,[1,3] Matthew M. Willmering, PhD,[1,2] Brandon Zanette, PhD,[4] Giles Santyr, PhD,[4,5] Sean B. Fain, PhD,[6,7,8] Michael J. Rock, MD,[9] Jaime Mata, PhD,[10] Deborah Froh, MD,[11] Sanja Stanojevic, PhD,[12] Zackary I. Cleveland, PhD,[1,2,3] Felix Ratjen, MD, PhD,[4,13] and Jason C. Woods, PhD[1,2]

Background: MRI with xenon-129 gas (Xe MRI) can assess airflow obstruction and heterogeneity in lung diseases. Specifically, Xe MRI may represent a sensitive modality for future therapeutic trials of cystic fibrosis (CF) therapies. The reproducibility of Xe MRI has not yet been assessed in the context of a multi-site study.
Purpose: To determine the same-day repeatability and 28-day reproducibility of Xe MRI in children with CF.
Study Type: Four-center prospective, longitudinal.
Population: Thirty-eight children (18 females, 47%), median interquartile range (IQR) age 12 (9–14) years old, with mild CF (forced expiratory volume in 1 second (FEV_1) ≥85% predicted).
Field Strength/Sequence: 3-T, two-dimensional (2D) gradient-echo (GRE) sequence.
Assessment: Xe MRI, FEV_1, and nitrogen multiple-breath wash-out for lung-clearance index ($LCI_{2.5}$) were performed. To assess same-day reproducibility, Xe MRI was performed twice within the first visit, and procedures were repeated at 28 days. Xe hypoventilation was quantified using ventilation-defect percentage (VDP) and reader-defect volume (RDV). For VDP, hypoventilated voxels from segmented images were identified using a threshold of <60% mean whole-lung signal and expressed as a percentage of the lung volume. For RDV, hypoventilation was identified by two trained readers and expressed as a percentage.
Statistical Tests: Inter-site comparisons were conducted using Kruskal–Wallis nonparametric tests with Dunn's multiple-comparisons tests. Differences for individuals were assessed using Wilcoxon matched-pairs tests. Bland–Altman tests were used to evaluate same-day repeatability, 28-day reproducibility, and inter-reader agreement. A *P*-value ≤0.05 was considered significant.
Results: Median FEV_1 %-predicted was 96.8% (86%–106%), and median $LCI_{2.5}$ was 6.6 (6.3–7.4). Xe MRI had high same-day reproducibility (mean VDP difference 0.12%, 95% limits of agreement [−3.2, 3.4]; mean RDV difference 0.42% [−2.5, 3.3]). At 28 days, 26/31 participants (84%) fell within the same-day 95% limits of agreement.
Data Conclusion: Xe MRI may offer excellent same-day and short-term reproducibility.
Evidence Level: 2
Technical Efficacy: Stage 2

J. MAGN. RESON. IMAGING 2024.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.29605

Received Jul 30, 2024, Accepted for publication Aug 27, 2024.

*Address reprint requests to: L.L.W., Cincinnati, OH, USA. E-mail: laura.walkup@cchmc.org

From the [1]Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA; [2]Department of Pediatrics, University of Cincinnati, Cincinnati, Ohio, USA; [3]Department of Biomedical Engineering, University of Cincinnati, Cincinnati, Ohio, USA; [4]Program in Translational Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; [5]Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada; [6]Department of Radiology, Carver College of Medicine, University of Iowa, Iowa City, Iowa, USA; [7]Department of Medicine, University of Wisconsin-Madison, Madison, Wisconsin, USA; [8]Department of Medical Physics, University of Wisconsin-Madison, Madison, Wisconsin, USA; [9]Department of Pediatrics, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA; [10]Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, Virginia, USA; [11]Department of Pediatrics, University of Virginia, Charlottesville, Virginia, USA; [12]Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada; and [13]Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada

Cystic fibrosis (CF) is a life-shortening, genetic condition affecting multiple organ systems.[1,2] Pulmonary manifestations of CF are primary factors for morbidity and mortality, as poor mucous clearance, recurrent inflammation, and infections lead to permanent lung structural remodeling with eventual, persistent lung-function decline.[1] Many studies have shown the connection between lung structural abnormalities and progressive lung-function changes, primarily described using the forced expiratory volume in 1 second (FEV_1) from spirometry.[1,2] These structural changes may manifest as early as infancy.[3]

With recent advancements and introduction of highly effective modulator therapies, such as the triple-combination elexacaftor/tezacaftor/ivacaftor (ETI) therapy that target specific classes of genetic mutations in the CF transmembrane conductance regulator (CFTR) protein, the landscape of CF lung disease has changed dramatically.[4,5] Many people with CF are living longer with improved quality of life and subclinical or mild lung disease.[6–8] Consequently, there is a need for improved, more sensitive endpoints for individualizing CF therapies and for new clinical trials in CF beyond the historical endpoints such as pulmonary exacerbations and changes in FEV_1.[9]

Ventilation imaging using inhaled hyperpolarized xenon-129 gas MRI (Xe MRI) is a safe and sensitive approach to measure regional ventilation deficits.[10] Studies in the CF population have shown Xe MRI has excellent sensitivity to subclinical airflow obstruction in children with normal FEV_1,[11] is robust to multiple quantitative analysis schemes,[12,13] is sensitive to treatment following a pulmonary exacerbation[14] and following airway-clearance therapy,[15] and has strong correlation with lung clearance index (LCI) from nitrogen multiple-breath washout.[16] These studies support Xe MRI as a sensitive outcome measure to quantify regional ventilation heterogeneity and treatment response in people with CF, especially those with sub-clinical or mild lung disease.[11–16] However, the same-day and short-term reproducibility of Xe MRI have not been assessed, particularly in the context of a multi-site study.

Against this background, this study aimed to assess whether Xe MRI is feasible in a multi-site study and has high same-day and 28-day reproducibility in children with mild CF lung disease.

Materials and Methods

Hyperpolarized Imaging for New Treatments (HyPOINT) is a prospective multi-site Xe MRI conducted at four sites in North America: Cincinnati Children's Hospital Medical Center (Site 1; Cincinnati, OH, USA), The Hospital for Sick Children (Site 2; Toronto, ON, Canada), The University of Virginia (Site 3; Charlottesville, VA, USA), and The University of Wisconsin (Site 4; Madison, WI, USA) under centralized Institutional Review Board approval (IRB 2019-1051) and US FDA IND (123577) for the three U.S.-based sites and for the Canadian site, under Research Ethics Board approval 1000065980 and Health Canada CTA 235275, with written informed consent for all participants or their parent/guardian.

Inclusion Criteria

Children with clinically stable CF were enrolled, and the inclusion criteria were age 6–18 years old at the time of informed consent, documentation of clinical CF diagnosis (defined as clinical features consistent with CF and either sweat chloride ≥60 mEq/L and/or two well-characterized pathogenic CFTR mutations), at least one F508del mutation, and FEV_1 %-predicted ≥80% per the Global Lung Function Initiative (GLI) reference equations. Clinical stability was defined as having no acute antibiotic use in the last 14 days prior to the first visit and no changes in chronic pulmonary medication or therapies (including CFTR modulator therapies) in the 28 days prior to the first visit. Exclusion criteria were standard MRI exclusions (eg, incompatible metal implants, claustrophobia), inability to cooperate with MRI procedures (eg, unable to complete a sufficient breath-hold maneuver), and pregnancy (i.e., positive urine pregnancy test).

Study Design and Procedures

This phase of the HyPOINT study involved two study visits separated by 28 days (±7 days), with MRI procedures performed twice during Visit 1 to assess same-day repeatability and once during Visit 2 to assess short-term 28-day reproducibility of Xe MRI. At each study visit, spirometry was collected per ATS/ERS standards,[17] and multiple-breath washout (MBW) was performed using Exhalyzer D equipment (EcoMedics AG, Duernten, Switzerland, operating with Spiroware software version 3.1.6) per published protocols to obtain the lung clearance index ($LCI_{2.5}$, the number of lung-volume turnovers to achieve 2.5% of the starting N_2 concentration) with external data analysis at the CFF MBW National Resource Center in Toronto.

All MRI acquisitions were conducted with 3 T scanners using either an Achieva (Philips Healthcare, Best, The Netherlands; Site 1), Prisma (Siemens Healthineers, Erlangen, Germany; Sites 2 and 3), or HealthCare Discovery scanner (General Electric Healthcare, Chicago, IL, USA; Site 4) with a harmonized gas dosing and acquisition protocol using a flexible vest-style Xe transmit-receive RF coil (Clinical MR Solutions, Brookfield, WI, USA). At each site, hyperpolarized Xe gas was prepared using a commercial polarizer (Model 9820 Polarean Inc., Durham, NC, USA) to approximately 25%–50% polarization, as measured by a local polarization measurement station (Polarean Inc., Durham, NC, USA). For ventilation imaging, the total administered gas volume was calculated as 1/6th of predicted total lung capacity (TLC) per plethysmography-based predictive equations,[18,19] up to 1 L maximum. The administered gas volume consisted of either 100% Xe gas at Site 1 or for the other three sites, Xe gas diluted with N_2 to achieve a total administered volume of ~1/6th TLC (eg, at Site 4, Xe gas was dosed at 10% of TLC, diluted with N_2 to achieve the total administered volume of 1/6th TLC). The gas was dispensed into a Tedlar bag (Jensen Inert Products, Coral Springs, FL, USA) for delivery to the participant. To account for different Xe polarizations and concentrations, a

Xe dose equivalence (DE) volume (i.e., the volume of 100% polarized, 100% ^{129}Xe-isotope enriched Xe) was calculated as follows:

$$\mathrm{DE} = V_{\mathrm{Xe}} \times P_{\mathrm{Xe}} \times f_{129_{\mathrm{Xe}}},$$

where V_{Xe} is the volume of Xe gas in mL, P_{Xe} is the fractional ^{129}Xe nuclear-spin polarization, and $f_{129_{\mathrm{Xe}}}$ is the ^{129}Xe-isotopic enrichment, which was typically 86% except for some instances at Site 2 where natural-abundance Xe gas was used (26% ^{129}Xe isotope) due to supply chain issues.

After initial screening, participants were positioned supine in the MRI scanner with appropriate hearing protection and pulse-oximetry monitoring. Baseline heart-rate and blood oxygen saturation (SpO_2) were recorded and monitored throughout the Xe MRI procedure to ensure participant safety, and Xe gas was administered in the presence of a medical professional, with vitals monitored for at least 2 minutes after each Xe dose to ensure recovery to baseline. After standard anatomical ^{1}H localization scans, Xe ventilation images were acquired during a coached inhalation and breath-hold maneuver (maximum duration 16 seconds) of hyperpolarized Xe gas. Coronal Xe ventilation images covering the whole-lung volume were acquired using the Xe MRI Clinical Trials Consortium consensus recommendations[20] consisting of a two-dimensional (2D) radiofrequency-spoiled gradient echo sequence with a resolution of $4 \times 4 \times 15$ mm^3 (reconstructed to a resolution of $2 \times 2 \times 15$ mm^3), with repetition time/echo time (TR/TE) = <10 msec/<5 msec, flip angle = 8°–12°, and scan duration up to 16 seconds, typically 8–12 seconds depending on the field of view and participant size.

For Visit 1, Xe ventilation images were acquired twice, separated by a minimum of 10 minutes where the participant was asked to stand up and walk after the first Xe ventilation images. The same dosing and acquisition strategy was used for only 1 acquisition at Visit 2, approximately 28 days after Visit 1, to assess short-term stability of Xe ventilation. The total time in the MRI scanner was approximately 40 minutes for Visit 1 and 30 minutes for Visit 2.

To facilitate Xe image segmentation and analysis, spatially matched anatomical ^{1}H images were acquired using the scanner's body coil with an inhalation and breath-hold maneuver of room air dosed and administered in a manner identical to the Xe gas (technical parameters: TR/TE = <10 msec/<5 msec; voxel size $4 \times 4 \times 15$ mm^3, reconstructed to a resolution of $2 \times 2 \times 15$ mm^3, with no slice gap, field of view same as Xe ventilation; flip angle 8°–12°; scan duration up to 16 seconds).

MRI Data Analysis

All MRI data were uploaded to a secure, online medical-imaging server (Ambra Health, Intelerad Medical Systems Inc., New York, NY, USA) for centralized data analysis conducted by the site in Cincinnati. The lungs were segmented manually (by DJR, 12 years of experience) from the Xe images using the anatomical ^{1}H images as guidance and regions of hypoventilation measured using custom software in MATLAB (The Mathworks Inc., Natick, MA, USA). Signal-to-noise ratio (SNR) for the Xe images was measured by dividing the mean signal within the segmented lung mask by the standard deviation of the background noise slice-by-slice, then all slices were averaged to determine the SNR for the whole Xe image set. Ventilation impairment on the Xe images was quantified in two ways: a ventilation defect percentage (VDP) and a reader-defect volume (RDV). In detail, VDP was calculated by first applying a non-parametric nonuniform intensity normalization (N4-ITK) bias-field correction algorithm[21] to the Xe images to correct B_1-inhomogeneity artifacts, then applying a threshold of <60% mean whole-lung Xe signal-intensity to define and detect ventilation deficits, with the VDP expressed as a percentage of the whole-lung volume in the Xe images per previous experience.[13] For the reader-defect analysis, two readers (DJR with 12 years of experience and JWP with 6 years of experience, both from Site 1) independently reviewed and manually selected regions of hypoventilation on the Xe images. These regions were then compared by the readers and a consensus RDV was calculated from the agreed-upon marked defects (hypoventilated regions which were marked by both readers) as a percentage of the whole-lung volume in the Xe images. Data sets with poor SNR (i.e., mean SNR <8 for the whole Xe image stack) were excluded from the VDP and RDV analysis based on previous work.[16]

Statistical Analysis

Statistical analysis was performed in GraphPad Prism (Version 10, GraphPad Software, Boston, MA, USA). Continuous variables were described using medians and interquartile ranges (IQRs), and categorical variables were described using percentages. Due to the distribution of the variables of interest, nonparametric tests were used. Inter-site comparisons were conducted using Kruskal–Wallis nonparametric tests with Dunn's multiple-comparisons tests. Furthermore, VDP and RDV differences for individual participants were compared using a two-sided, Wilcoxon matched-pairs test. The same-day repeatability of VDP and RDV within the same participant was described using the mean difference and 95% limits of agreement from Bland–Altman plots. Bland–Altman tests were used to evaluate the 28-day reproducibility of VDP and consensus RDV measurements. Scatter plots were used to assess relationships between Xe MRI outcomes and PFTs, specifically FEV_1 %-predicted and $LCI_{2.5}$. Bland–Altman analysis also was used to evaluate inter-rater agreement between individual reader RDV values using data from both visits (a total of 102 image sets). A P-value ≤0.05 was considered statistically significant.

Results

Table 1 shows demographics and clinical characteristics of the 38 children with CF at Visit 1. The median participant age at Visit 1 was 12 years (IQR 9–14), and subjects at Site 2 were older than those at Site 4 with median ages of 14 years (IQR 12–15) and 9 years (IQR 7–11), respectively. There were no other significant differences in participant age across the sites ($P = 0.77$ Site 1 vs. Site 2; $P = 1.0$ Site 1 vs. Site 3; $P = 0.87$ Site 1 vs. Site 4; $P = 1.0$ Site 2 vs. Site 3; $P = 0.20$ Site 3 vs. Site 4). The median FEV_1 %-predicted at Visit 1 was 96.6% (IQR 86%–104%), with no significant inter-site differences ($P = 0.83$ Site 1 vs. Site 2; $P = 1.0$ Site 1 vs. Site 3; $P = 1.0$ Site 1 vs. Site 4; $P = 0.055$ Site 2 vs. Site 3; $P = 0.51$ Site 2 vs. Site 4; $P = 1.0$ Site 3 vs. Site 4). The median $LCI_{2.5}$ was 6.6 (IQR 6.3–7.4) with a lower LCI

TABLE 1. Subject Demographics, Clinical Characteristics, and Xe MRI at Visit 1

	All Participants (N = 38)	Site 1 (N = 11)	Site 2 (N = 10)	Site 3 (N = 8)	Site 4 (N = 9)
Age, years	12 (9–14)	11 (9–14)	14 (12–15)	13 (11–14)	9 (7–11)
Female sex, N (%)	18 (47)	4 (36)	5 (50)	6 (75)	3 (33)
BMI, kg/m^2	19.2 (15.5–20.8)	19.0 (17.8–22.1)	19.5 (18.5–20.8)	20.9 (18.7–21.9)	18.2 (16.4–19.2)
FEV_1, %-predicted	96.7 (86–106)	96.6 (85.4–101.0)	92.5 (83.8–96.3)	102.1 (86.7–113.0)	103.5 (95.0–109.3)
$LCI_{2.5}$	6.6 (6.3–7.4)	6.2 (5.8–6.6)	7.3 (6.5–9.8)	6.3 (6.3–7.5)	6.6 (6.3–7.3)
VDP, % (Visit 1, Scan 1)	5.0 (2.8–7.6)	4.1 (2.8–6.8)	7.5 (4.9–11.6)	4.0 (1.3–7.0)	4.1 (3.6–5.6)[a]
RDV, % (Visit 1, Scan 1)	1.2 (0.0–3.1)	0.4 (0.0–1.5)	5.4 (1.4–11.5)	0.77 (0.0–2.7)	0.5 (0.0–2.2)[a]

Values are reported as median (IQR) unless indicated otherwise. Inter-site comparisons were conducted using a Kruskal–Wallis nonparametric test with Dunn's multiple-comparisons test, with $P \leq 0.05$ considered statistically significant.
BMI = body mass index; FEV_1 = forced expiratory volume in 1 second; $LCI_{2.5}$ = lung clearance index (i.e., the number of lung-volume turnovers to achieve 2.5% of the starting N_2 concentration); VDP = ventilation defect percentage; RDV = reader defect volume.
[a]Excluding three data sets with low signal-to-noise ratio (SNR <8).

at Site 1 compared with Site 2, median $LCI_{2.5}$ values of 6.2 (IQR 5.8–6.6) and 7.3 (IQR 6.5–9.8), respectively. There were no other significant differences in LCI across the sites ($P = 0.60$ Site 1 vs. Site 3; $P = 0.85$ Site 1 vs. Site 4; $P = 1.0$ Site 2 vs. Site 3, $P = 1.0$ Site 2 vs. Site 4; $P = 1.0$ Site 3 vs. Site 4).

The Xe MRI procedure was well tolerated by all participants. For Visit 1 across all sites, the median Xe dose-equivalent volume administered was 77 mL (IQR 49–176), and the median Xe MRI SNR was 35.4 (IQR 17.6–44.8) for all Visit 1 Xe image sets (i.e., including both Scan 1 and Scan 2). Figure 1 shows an inter-site comparison of Xe dose-equivalent volume and SNR. Site 1 administered a higher Xe dose-equivalent volume, a median of 194 mL (IQR 179–220), compared to the other three sites (Fig. 1a). There were no other significant differences in DE between sites ($P = 1.0$ Site 2 vs. Site 3; $P = 0.22$ Site 2 vs. Site 4; $P = 1.0$ Site 3 vs. Site 4). In Fig. 1b, Site 4 had significantly lower Xe SNR, a median of 9.5 (IQR 8.1–11.7), compared to the other three sites but there were no other significant differences in SNR between sites ($P = 0.15$ Site 1 vs. Site 2, $P = 1.0$ Site 1 vs. Site 3; $P = 0.54$ Site 2 vs. Site 3). Three Xe MRI data sets, all Visit 1 Scan 1 data from Site 4, had a median SNR of ~6.5 and were excluded from subsequent VDP and RDV analysis due to poor SNR (<8). Across all sites, median VDP at Visit 1 Scan 1 was 5.0% (IQR 2.8%–7.6%), and there were no significant differences across sites ($P = 0.71$ Site 1 vs. Site 2; $P = 1.0$ Site 1 vs. Site 3; $P = 1.0$ Site 1 vs. Site 4; $P = 0.49$ Site 2 vs. Site 3; $P = 1.0$ Site 2 vs. Site 4; $P = 1.0$ Site 3 vs. Site 4). The median RDV for Visit 1, Scan 1 across all sites was 1.2% (IQR 0.0%–3.1%), and there were no significant differences across sites ($P = 0.24$ Site 1 vs. Site 2; $P = 1.0$ Site 1 vs. Site 3; $P = 1.0$ Site 1 vs. Site 4; $P = 0.81$ Site 2 vs. Site 3; $P = 0.26$ Site 2 vs. Site 4; $P = 1.0$ Site 3 vs. Site 4). Figure 2 shows representative images from a participant at each site.

In Fig. 3, the same-day repeatability of Xe MRI is shown. The median time between Scan 1 and Scan 2 was 33 minutes (IQR 10–87 minutes). Figure 3a,c show the differences in same-day VDP and RDV measurements, respectively, for individual participants with no significant changes ($P = 0.34$ for VDP; $P = 0.16$ for RDV), and there was no pattern between the number of minutes between scans and the differences in the VDP and RDV measurements, i.e., participants who had longer time between scans did not appear to have larger Xe MRI changes compared to those with a smaller time difference. The mean difference of VDP was 0.12% ± 1.7% with 95% limits of agreement of −3.2% to 3.4% (Fig. 3b). For the consensus RDV measurements in Fig. 3d, the mean difference was 0.42% ± 1.5%, with 95% limits of agreement of −2.5% to 3.3%. The same-day difference in VDP and RDV measurements were plotted against FEV_1 %-predicted and $LCI_{2.5}$ in Fig. 4. Differences in VDP and RDV between repeated measures in the same day were not explained by baseline pulmonary function.

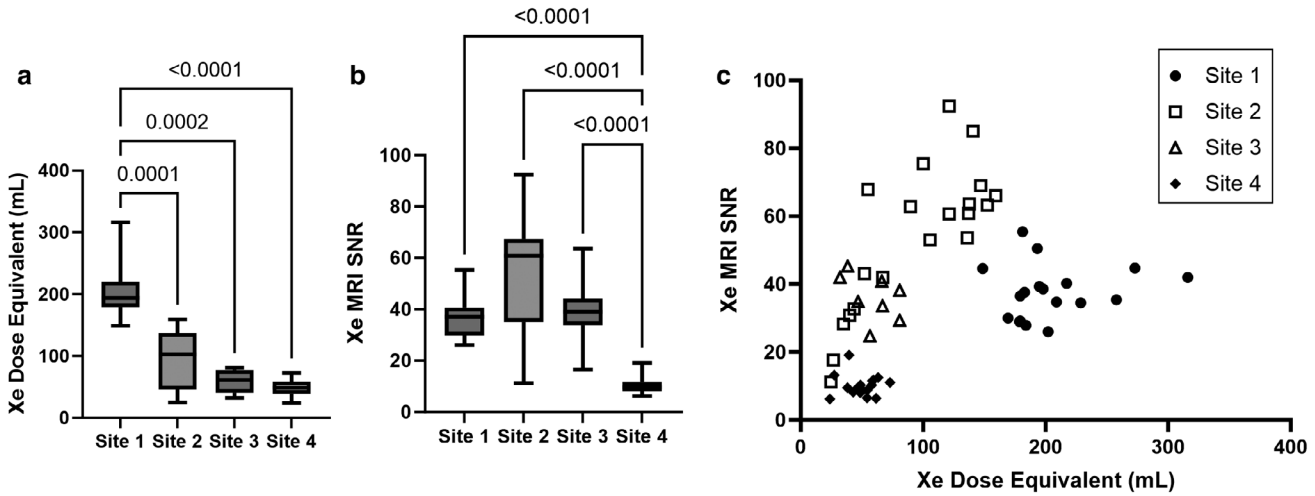

FIGURE 1: Inter-site comparison of Xe dose equivalent (DE) (a) and Xe MRI signal-to-noise ratio (SNR) (b). Box-and-whisker plots, denoting the minimum, 25th percentile, median, 75th percentile, and maximum are shown, with inter-site comparisons conducted using Kruskal–Wallis nonparametric tests with Dunn's multiple-comparisons tests. Panel (c) is a scatter plot of Xe MRI SNR vs. Xe dose-equivalent volume with markers differentiating the study sites (64 instances from Visit 1 data).

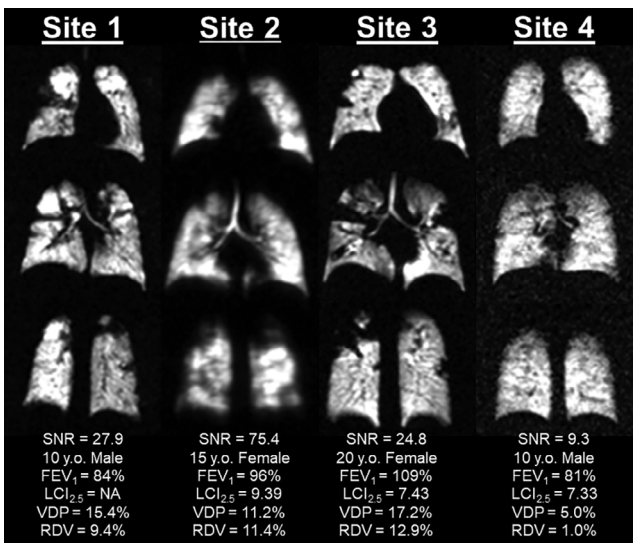

FIGURE 2: Inter-site comparison of Xe image quality. Four representative cases with demographics, FEV_1, and $LCI_{2.5}$ are shown, one from each site. These were all data from Visit 1, Scan 1 of the study. Xe MRI SNR, VDP, and consensus RDV measurements are given.

To assess 28-day reproducibility of Xe ventilation MRI, Visit 2 occurred at a median of 28 days (IQR 27–34 days) after the Visit 1 baseline. Two participants were excluded from analysis: one due to the onset of sinusitis with productive cough and weight loss, with 18% absolute decline in FEV_1 %-predicted between visits, and a second participant was excluded for large FEV_1 %-predicted decline (14% absolute). Figure 5 shows the individual trajectories of VDP and RDV, with no significant change ($P = 0.65$ for VDP; $P = 0.22$ for RDV). Figure 5 also shows the 28-day change in Xe MRI outcomes compared to the 28-day relative percentage change in FEV_1 %-predicted and $LCI_{2.5}$. Most of the participants (84%; 26/31 participants) fall within the same-day limits of agreement for the Xe MRI outcomes. While most of the participants fell within the same-day limits of agreement, there were a few outliers. Upon review of the interval clinical history for these outliers, it was determined one participant had clinically noted adherence issues, another had reported an ongoing cough, and another participant had a new onset infection, thus the variation in Xe MRI observed likely was reflective of real physiological variation in these individuals.

The difference between the two readers was −0.2% with 95% limits of agreement −2.9% to 3.2% (Fig. 6a) Furthermore, as shown in Fig. 6b, the difference between RDV and VDP was 3.2% with 95% limits of agreement −0.7% to 7.0%, with VDP systematically producing a higher measurement of ventilation heterogeneity compared to RDV.

Figure 7 demonstrates the same-day repeatability and 28-day reproducibility of Xe MRI in three representative participants: one with little or no change across all three scans, which was representative of most participants; another participant with large same-day ventilation changes and a third with good same-day reproducibility but ventilation changes at Visit 2. Representative images are shown in the middle column of Fig. 7 (subject CF2), where there was a small improvement in ventilation in the apices of the lungs seen in Scan 2 of Visit 1. This subtle ventilation change likely was related to the movement of a transient airflow-obstructive feature like a mucus plug in the brief time interval between scans. At Visit 2, the ventilation pattern appeared more similar to that of Scan 1, Visit 1. Conserved ventilation deficits, i.e., regions of hypoventilation that are observed consistently across time, likely were reflective of stable and unchanged airflow-obstruction features, such as airway remodeling or

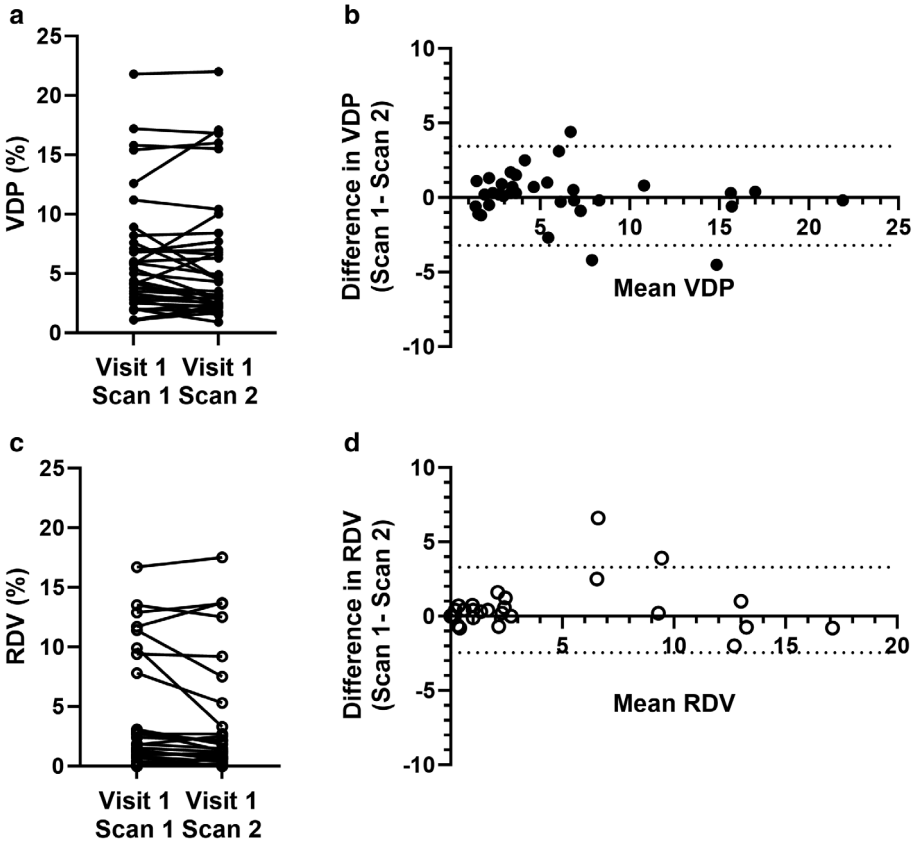

FIGURE 3: Same-day repeatability of Xe MRI VDP and consensus RDV measurements. (a) Individual VDP changes are shown between Scan 1 and Scan 2, and in the Bland–Altman plot in (b), the mean difference in VDP measurements was 0.12% with 95% limits of agreement −3.2% to 3.4%. For the consensus RDV measurement, individual changes between Scan 1 and Scan 2 are shown in (c) and in (d) the mean RDV difference was 0.42% with 95% limits of agreement of −2.5% to 3.3%.

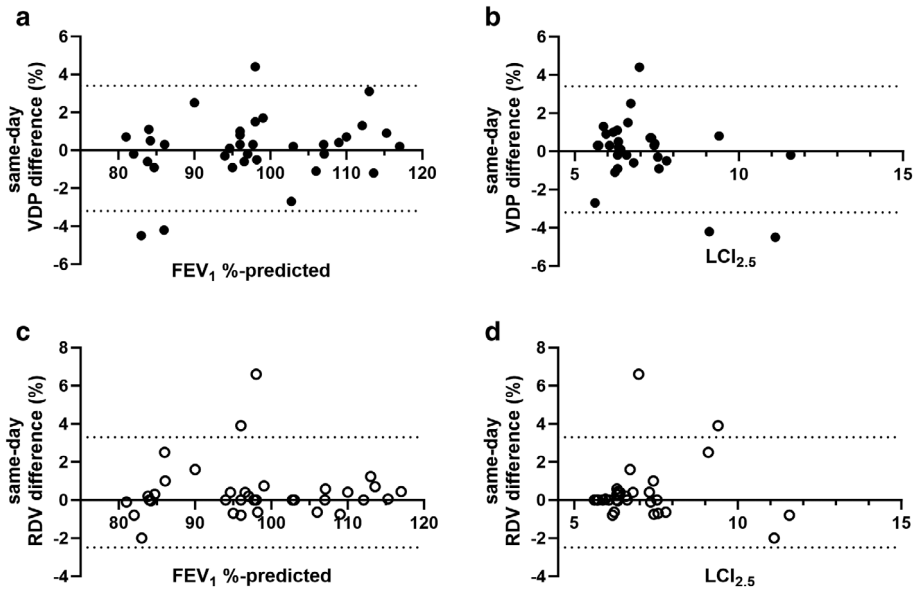

FIGURE 4: Same-day repeatability of Xe MRI vs. PFTs. In (a) and (b), the same-day difference in VDP is compared to FEV_1 %-predicted and $LCI_{2.5}$, respectively. In (c) and (d), the same comparisons are made for RDV. The dashed lines on the panels are the same-day limits of agreement for VDP (−3.2% to 3.4%) and RDV (−2.5% to 3.3%). There was no relationship between pulmonary function and variability in the same-day Xe MRI measurements.

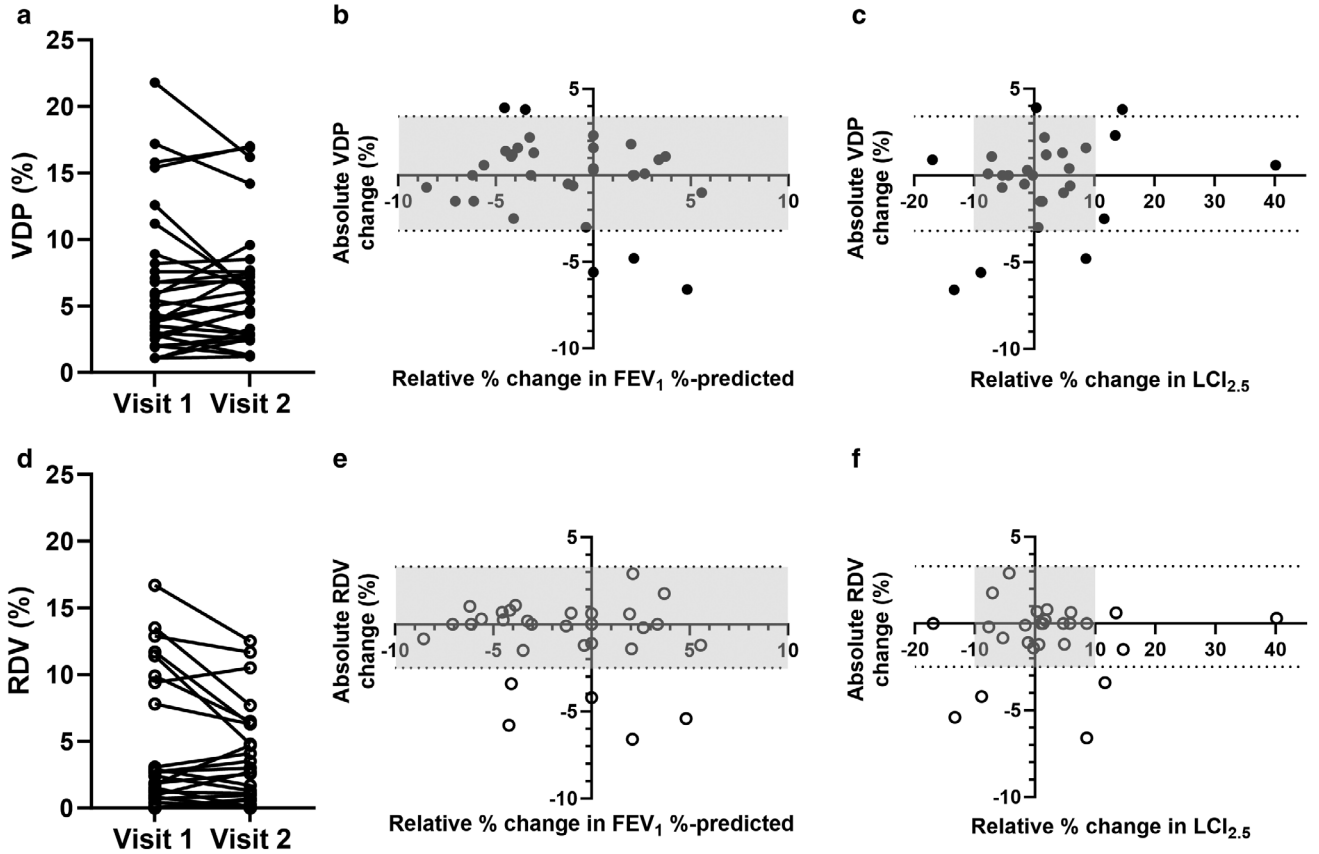

FIGURE 5: The 28-day reproducibility of Xe MRI. In (a) and (d), individual trajectories for VDP and RDV are shown, respectfully, with no statistically significant change. In (b) and (c), the 28-day change in VDP is plotted against the relative 28-day changes in FEV_1 and $LCI_{2.5}$, respectively, with the dashed lines representing the same-day limits of agreement for VDP (−3.2% to 3.4%). In (e) and (f), the 28-day change in RDV is compared to relative 28-day changes in FEV_1 and $LCI_{2.5}$ with the dashed lines representing the same-day limits of agreement for RDV (−2.2% to 3.3%). The gray boxes are bound by the same-day Xe MRI limits of agreement and by ±10% relative change in PFTs.

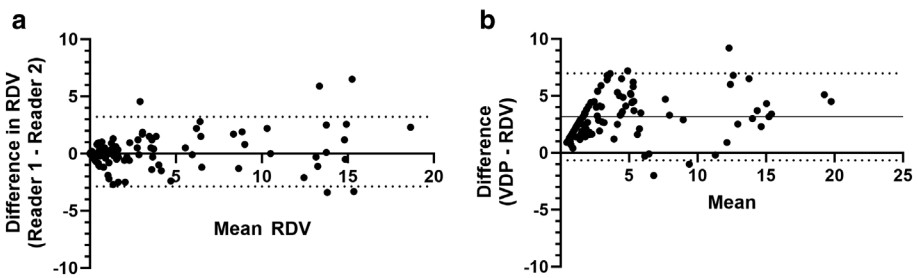

FIGURE 6: (a) Bland–Altman analysis to assess inter-rater agreement in RDV measurements. The mean difference in RDV between the two readers was 0.2% with 95% limits of agreement −2.9% to 3.2%. (b) Bland–Altman analysis to compare VDP and consensus RDV measurements. The mean difference was 3.2% with 95% limits of agreement −0.7% to 7.0%, with VDP systematically producing a higher measurement of ventilation heterogeneity compared to the consensus RDV. For both these panels, Xe scans from Visit 1 (two scans per participant) and Visit 2 were combined.

bronchiectasis which would impact ventilation to the distal lung. Subject CF3 (Fig. 7, right column) had ventilation improvement between Visit 1 and Visit 2, with reductions in both VDP and RDV; the participant had consistent 82% FEV_1 between visits but an improvement in $LCI_{2.5}$ (11.6 at Visit 1, 10.5 at Visit 2), which may suggest Xe MRI like MBW is more sensitive to changes in ventilation heterogeneity than FEV_1.

Discussion

In this study, we demonstrated how a harmonized Xe MRI ventilation acquisition on three different MRI scanner platforms may have high potential for clinical translation and for application in larger multi-site trials using Xe MRI. Over the last 20 years, multiple single-site studies have supported the potential of Xe MRI as a sensitive assessment of regional

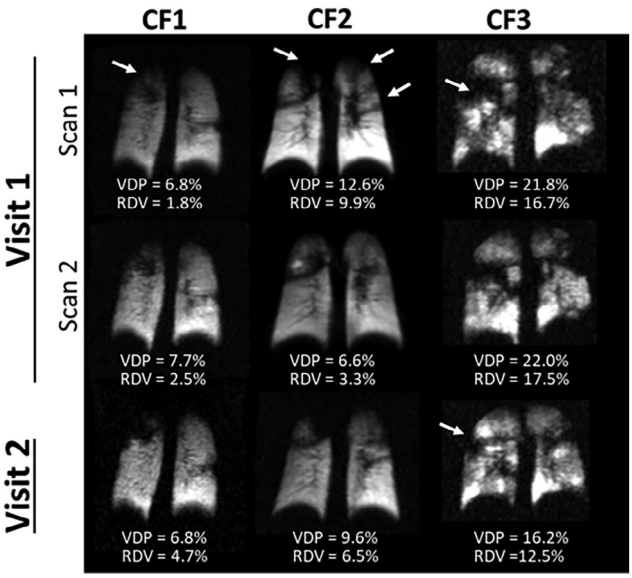

FIGURE 7: Representative single coronal Xe ventilation MRI images from Visit 1 (top two rows for Scan 1 and Scan 2) and from Visit 2 (bottom row) for three participants. VDP and RDV measurements are shown below the images. CF1 (left column) showed stable ventilation with conserved ventilation deficits appearing in the same location across all three scans, and similar VDP and RDV measurements. This stability of ventilation pattern was typical of most participants. CF2 (middle column) had large same-day ventilation changes during Visit 1, where the two scans were separated by 19 minutes; the arrows highlight ventilation deficits that were present at the first scan but not in the second scan. Notice the ventilation pattern at Visit 2 looks similar to that of Visit 1, Scan 1. For CF3 (right column), same-day ventilation in Visit 1 was consistent, but there was an improvement in ventilation at Visit 2 with a reduction in both VDP and RDV and improvement in the right-lung ventilation deficit marked by arrows. This participant had consistent 82% FEV_1 at both visits but improvement in LCI2.5 (11.573 at Visit 1, 10.542 at Visit 2).

ventilation deficits in mild CF lung disease, yet critically, no studies to date have shown the same-day reproducibility ventilation deficits from Xe MRI in the context of a prospective multi-site study. With the recent FDA approval of Xe MRI to assess regional lung ventilation in people ages 12 and older, more institutions are likely to adopt Xe MRI in their clinical practice and as outcomes in clinical trials, and while there is strong potential for clinical utility in CF, understanding the variability of Xe MRI is essential.

The inter-site differences in Xe dose-equivalent volume can be attributed to differences in Xe hyperpolarizer performance coupled with differences in amount of Xe gas in the dose administered. The protocol allowed flexibility to administer either 100% Xe gas, as was the case at Site 1, or a Xe/N_2 mixture, as was done at the three other sites, permitting that the total administered gas volume was 1/6th predicted TLC and the image SNR was sufficiently high (>8) to allow for quantification. However, in the context of a multi-site study, doubling the dose-equivalent Xe volume does not necessarily guarantee double the image SNR; for instance, Site 2 administered approximately half the dose-equivalent Xe volume as Site 1 yet had higher SNR than Site 1. Three Xe MRI datasets, all Visit 1 Scan 1 data from Site 4 which occurred earlier in the study, were excluded from subsequent VDP and RDV analysis due to poor SNR (<8). The small SNR difference likely was reflective of Xe dosing preferences at that site, which was adjusted in subsequent study visits. While there are more sophisticated, non-Cartesian pulse sequences for Xe ventilation,[22,23] this study used a standard Cartesian GRE sequence with acquisition parameters recommended by the Xe MRI Clinical Trials Consortium consensus guidelines to limit variability across sites. The inter-site differences in SNR are complex and likely related to site- and scanner-specific hardware and parameters, like coil placement, noise profile, and MRI-scanner-vendor image reconstruction. However, despite these inter-site differences in dosing strategy, hardware configurations, and software settings, the same extent of hypoventilation was quantified, with no significant differences in VDP or RDV between sites, and high inter-visit reproducibility was observed across this multi-center cohort of patients with CF and high FEV_1 %-predicted. This agrees with a two-site, retrospective study of children with mild CF reported by Couch et al that showed similar VDP measurements between two sites with different MRI scanners, coil configurations, and dosing strategies.[12]

In this study, ventilation deficits were quantified using a computerized threshold-based cutoff technique (VDP) and a manual reader-scoring method (RDV), both of which were robust to inter-site differences in acquisition and dosing. In our study, the same-day mean difference in VDP was $0.12\% \pm 1.7\%$. The same-day reproducibility supports that the ventilation deficits in mild CF lung disease observed with Xe MRI are stable, and that the VDP and RDV measurements from Xe MRI are reproducible over a short period of time. This agrees with Kanhere et al who showed no significant difference by a paired *t*-test of same-day repeated VDP measurements in a single-center study of a small group of children with CF (mean VDP $7.2\% \pm 3.4\%$ with a range of 1.8%–13.0%) and healthy controls (mean VDP $1.8\% \pm 0.8\%$, range 0.9%–2.9%).[16] Svenningsen and colleagues showed highly reproducible Xe MRI VDP measurements in a cohort of seven adults with severe asthma each of whom were imaged at two geographic sites within 24 hours, showing high inter-site reproducibility in physiological location and extent of ventilation deficits in the same participants.[24] They reported a 3% mean VDP difference and 95% limits of agreement of −14% to 8%, which is more variable than the same-day variation observed in this study; this may be due to differences in MRI acquisitions including dosing, timing of the MRI exams, and VDP quantification and perhaps also related to pathophysiological differences between these diseases. Ebner et al showed highly reproducible same-day VDP

measurements in a group of healthy adults and those with asthma and reported a mean difference of -0.88 ± 1.52, which was largely attributed to two participants with a ~4% VDP difference between scans.[25] Likewise, Smith and colleagues showed a bias of 0.2% in same-day repeated VDP measurements in a cohort of people with CF across a range of disease severity.[26] A study by Kirby et al using helium-3 MRI assessed the within-subject, between-test variability of the VDP and reported limits of reproducibility (1.96 × the standard deviation) less than 4% and that the smallest detectable difference in VDP was 2%.[27] In total, this supports that the location and extent of ventilation deficits observed with Xe MRI are highly reproducible over a short period of time.

At Visit 2, VDP and RDV were stable with minimal within-participant changes, supporting that Xe MRI had excellent short-term stability in clinically stable people with CF. A longitudinal study by Smith et al of 29 people with moderate CF lung disease (mean FEV_1 %-predicted 71% ± 26%) showed no significant change in Xe MRI at a median follow-up interval of 16 months, but there was large inter-subject variability in the extent and direction of Xe ventilation changes independent of disease severity.[26] Alam and colleagues recently demonstrated same-day and 1-month reproducibility of dynamic, multiple-breath washout Xe MRI in a single-center cohort of healthy and CF children,[28] which agrees with our results and supports the stability of ventilation heterogeneity and Xe MRI in people with mild CF lung disease over a short period of time. Determining the intrinsic same-day variability of Xe MRI and the threshold of clinically-relevant changes in VDP remain critical aspects for the clinical translation of Xe MRI. A single-site study of Xe MRI in CF lung disease by Smith et al defined ±1.6% VDP as a relevant change, as defined by the 95% limits of agreement from same-day VDP measurements.[26] Applying this metric to this multi-site study, the threshold for relevant VDP and RDV change would be ±3.5% and ±1.9%, respectively. The relatively higher VDP threshold in this study likely is related to differences in VDP definition and measurement between studies and further compounded by factors intrinsic to the multi-center nature of this study including the variability in dosing and acquisition technique, scanner hardware, image reconstruction, and coil-placement differences.

Regarding the determination of a clinically-relevant difference in VDP, one study reported a mean 3% VDP improvement in a group of children with CF treated with antibiotics for pulmonary exacerbation.[14] Another study of hyperpolarized helium-3 gas ventilation MRI in a cohort of adults with asthma reported a minimally clinically important difference in VDP of 4% as anchored to the patient-reported and clinically-validated Asthma Control Questionnaire (AQL).[29] Despite the differences in patient populations, data acquisition, and study goals, these studies support that a VDP change of ~2%–3% likely is clinically relevant and raises important considerations for gas dosing, image acquisition, and VDP analysis in future multi-site clinical trials using outcomes from Xe MRI. While this study was designed to assess Xe-MRI reproducibility over the course of 1 month in children with CF, there are important age- and sex-related changes in Xe MRI which would need to be considered the design of longer longitudinal studies and in larger clinical trials with Xe MRI as an outcome.[30–33] Furthermore, a study by Garrison et al showed reproducible same-day Xe gas-exchange MRI metrics related to membrane uptake and RBC transfer in a single-center study of healthy controls and those with chronic obstructive pulmonary disease (COPD) and also showed variations in these metrics depending on lung inflation, which again emphasizes the importance of gas dosing in future Xe MRI studies.[34]

While previous studies like[14,29] used changes in clinical parameters like symptoms or questionnaires to define change in Xe MRI, in this study, we have used a distributional approach to define reproducibility. With that said, regional changes are often obvious and quantifiable, leaving the potential for much higher sensitivity to true physiological change. Understanding Xe MRI change in the background of stable CF lung disease will be important to determining treatment response, and Xe MRI showed excellent 28-day stability with no significant changes in VDP or RDV.

Limitations

The number of participants at each site was relatively small, and thus some of the statistical comparisons may be underpowered, and inter-reader reliability was not assessed as part of this study. As this study focused on mild CF lung disease, repeatability and stability of Xe MRI may be different in those with advanced CF lung disease or other patient populations where the disease pathophysiology of airflow obstruction is different. The flexibility in Xe dosing (i.e., total administered volume of up to 1/6th predicted TLC, composed of either 100% Xe or a mixture of Xe/N_2) was intended to allow individual sites to operate within their current operating standards for research Xe MRI studies. Dosing strategy likely impacts the distribution of gas within the lungs, and in the future, specification of Xe dose-equivalent volume may help minimize some of the minor inter-site variation in SNR seen in this study. Furthermore, Xe MRI is a relatively niche modality considering the multi-nuclear hardware, hyperpolarized, and expertise needed, which may limit its accessibility in future clinical trials and in clinical practice; however, the number of sites performing Xe MRI continues to grow. While outside of the scope of the current work, it is important to recognize there are other MRI-based techniques for assessing regional lung ventilation including three-dimensional (3D) phase-resolved functional lung (PREFUL)[35,36] and matrix-pencil/Fourier decomposition methods[37,38] which do not require an inhaled contrast agent.

Conclusion

Using a multi-site, harmonized Xe gas dosing and acquisition protocol, Xe MRI may show to be highly repeatable within the same day and could show excellent short-term stability in a cohort of children with stable, mild CF lung disease.

Acknowledgments

The authors would like to acknowledge the HyPOINT study-team members at Cincinnati Children's Hospital (Kelly Thornton, Carrie Stevens, Megan Schmitt, and Carter McMaster); Hospital for Sick Children Toronto (Aviva West, Jacky Au, Sharon Braganza, Daniel Li, Samal Munidasa, and Yonni Friedlander); University of Virginia (Roselove Asare and Rene Herrera); and University of Wisconsin-Madison for their contributions to this study. Data from this study were presented in part at the American Thoracic Society ATS 2021 and 2022 meetings and the North American Cystic Fibrosis Conference (NACFC) 2021. This work was supported by the Cystic Fibrosis Foundation (WOODS19A0) and University of Cincinnati Center for Clinical and Translational Science and Training (National Institutes of Health Clinical and Translational Science Award UL1TR001425).

References

1. Mall MA, Burgel P-R, Castellani C, Davies JC, Salathe M, Taylor-Cousar JL. Cystic fibrosis. Nat Rev Dis Primers 2024;10(1):53.
2. Swetland DV, Savant AP. Cystic fibrosis year in review 2023. Pediatr Pulmonol 2024. https://doi.org/10.1002/ppul.27190
3. Roach DJ, Cremillieux Y, Fleck RJ, et al. Ultrashort echo-time magnetic resonance imaging is a sensitive method for the evaluation of early cystic fibrosis lung disease. Ann Am Thorac Soc 2016;13(11):1923-1931.
4. Stahl M, Dohna M, Graeber SY, et al. Impact of elexacaftor/tezacaftor/ivacaftor therapy on lung clearance index and magnetic resonance imaging in children with cystic fibrosis and one or two F508del alleles. Eur Respir J 2024;64:2400004.
5. David M, Benlala I, Bui S, et al. Longitudinal evaluation of bronchial changes in cystic fibrosis patients undergoing elexacaftor/tezacaftor/ivacaftor therapy using lung MRI with ultrashort echo-times. J Magn Reson Imaging 2024;60(1):116-124.
6. Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry 2022 Annual Data Report. Bethesda, Maryland2022.
7. Kerem E, Orenti A, Adamoli A, Hatziagorou E, Naehrlich L, Sermet-Gaudelus I. Cystic fibrosis in Europe: Improved lung function & longevity: Reasons for cautious optimism, but challenges remain. Eur Respir J 2024;63:2301241.
8. Sutharsan S, Dillenhoefer S, Welsner M, et al. Impact of elexacaftor/tezacaftor/ivacaftor on lung function, nutritional status, pulmonary exacerbation frequency and sweat chloride in people with cystic fibrosis: Real-world evidence from the German CF Registry. Lancet Reg Health Eur 2023;32:100690.
9. Bayfield KJ, Weinheimer O, Middleton A, et al. Comparative sensitivity of early cystic fibrosis lung disease detection tools in school aged children. J Cyst Fibros 2024. https://doi.org/10.1016/j.jcf.2024.05.012
10. Stewart NJ, Smith LJ, Chan HF, et al. Lung MRI with hyperpolarised gases: Current & future clinical perspectives. Br J Radiol 2022;95(1132):20210207.
11. Thomen RP, Walkup LL, Roach DJ, Cleveland ZI, Clancy JP, Woods JC. Hyperpolarized 129Xe for investigation of mild cystic fibrosis lung disease in pediatric patients. J Cyst Fibros 2017;16(2):275-282.
12. 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.
13. Roach DJ, Willmering MM, Plummer JW, et al. Hyperpolarized (129)xenon MRI ventilation defect quantification via thresholding and linear binning in multiple pulmonary diseases. Acad Radiol 2021;29:S145–S155.
14. Rayment JH, Couch MJ, McDonald N, et al. Hyperpolarised (129)Xe MRI to monitor treatment response in children with cystic fibrosis. Eur Respir J 2019;53:1802188.
15. West ME, Spielberg DR, Roach DJ, et al. Short-term structural and functional changes after airway clearance therapy in cystic fibrosis. J Cyst Fibros 2023;22:926-932.
16. 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(8):1073-1075.
17. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of spirometry 2019 update. An Official American Thoracic Society and European Respiratory Society Technical Statement. Am J Respir Crit Care Med 2019;200(8):e70-e88.
18. Stocks J, Quanjer PH. Reference values for residual volume, functional residual capacity and total lung capacity. ATS Workshop on Lung Volume Measurements. Official Statement of The European Respiratory Society. Eur Respir J 1995;8(3):492-506.
19. Zapletal A, Paul T, Samanek M. Significance of contemporary methods of lung function testing for the detection of airway obstruction in children and adolescents (author's transl). Z Erkr Atmungsorgane 1977;149(3):343-371.
20. Niedbalski PJ, Hall CS, Castro M, et al. Protocols for multi-site trials using hyperpolarized 129Xe MRI for imaging of ventilation, alveolar-airspace size, and gas exchange: A position paper from the 129Xe MRI clinical trials consortium. Magn Reson Med 2021;86(6):2966-2986.
21. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 bias correction. IEEE Trans Med Imaging 2010;29(6):1310-1320.
22. Fang Y, Li H, Shen L, et al. Rapid pulmonary 129Xe ventilation MRI of discharged COVID-19 patients with zigzag sampling. Magn Reson Med 2024;92(3):956-966.
23. Plummer JW, Hussain R, Bdaiwi AS, et al. A decay-modeled compressed sensing reconstruction approach for non-Cartesian hyperpolarized (129)Xe MRI. Magn Reson Med 2024;92(4):1363-1375.
24. Svenningsen S, McIntosh M, Ouriadov A, et al. Reproducibility of hyperpolarized 129Xe MRI ventilation defect percent in severe asthma to evaluate clinical trial feasibility. Acad Radiol 2021;28(6):817-826.
25. Ebner L, He M, Virgincar RS, et al. Hyperpolarized 129Xenon magnetic resonance imaging to quantify regional ventilation differences in mild to moderate asthma: A prospective comparison between semiautomated ventilation defect percentage calculation and pulmonary function tests. Invest Radiol 2017;52(2):120-127.
26. Smith LJ, Horsley A, Bray J, et al. The assessment of short and long term changes in lung function in CF using (129)Xe MRI. Eur Respir J 2020;56:2000441.
27. 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(8):1006-1013.
28. Alam FS, Zanette B, Munidasa S, et al. Intra- and inter-visit repeatability of (129) xenon multiple-breath washout MRI in children with stable cystic fibrosis lung disease. J Magn Reson Imaging 2023;58(3):936-948.
29. Eddy RL, Svenningsen S, McCormack DG, Parraga G. What is the minimal clinically important difference for helium-3 magnetic resonance imaging ventilation defects? Eur Respir J 2018;51(6):1800324.

30. Collier GJ, Smith LJ, Saunders LC, et al. Age, sex, and lung volume dependence of dissolved xenon-129 MRI gas exchange metrics. Magn Reson Med 2024;92(4):1471-1483.

31. Plummer JW, Willmering MM, Cleveland ZI, Towe C, Woods JC, Walkup LL. Childhood to adulthood: Accounting for age dependence in healthy-reference distributions in (129) Xe gas-exchange MRI. Magn Reson Med 2023;89(3):1117-1133.

32. Mummy D, Zhang S, Bechtel A, et al. Functional gas exchange measures on (129)Xe MRI and spectroscopy are associated with age, sex, and BMI in healthy subjects. Front Med (Lausanne) 2024;11:1342499.

33. Rao Q, Li H, Zhou Q, et al. Assessment of pulmonary physiological changes caused by aging, cigarette smoking, and COPD with hyperpolarized (129)Xe magnetic resonance. Eur Radiol 2024. https://doi.org/10.1007/s00330-024-10800-w

34. Garrison WJ, Qing K, He M, et al. Lung volume dependence and repeatability of hyperpolarized (129)Xe MRI gas uptake metrics in healthy volunteers and participants with COPD. Radiol Cardiothorac Imaging 2023;5(3):e220096.

35. Dohna M, Voskrebenzev A, Klimeš F, et al. PREFUL MRI for monitoring perfusion and ventilation changes after elexacaftor-tezacaftor-ivacaftor therapy for cystic fibrosis: A feasibility study. Radiol Cardiothorac Imaging 2024;6(2):e230104.

36. Friedlander Y, Munidasa S, Thakar A, et al. Phase-resolved functional lung (PREFUL) MRI to quantify ventilation: Feasibility and physiological relevance in severe asthma. Acad Radiol 2024;31:3416-3426.

37. Bauman G, Bieri O. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. Magn Reson Med 2017;77(1):336-342.

38. Doellinger F, Bauman G, Roehmel J, et al. Contrast agent-free functional magnetic resonance imaging with matrix pencil decomposition to quantify abnormalities in lung perfusion and ventilation in patients with cystic fibrosis. Front Med (Lausanne) 2024;11:1349466.