Radiology

Exercise-induced Bronchoconstriction: Reproducibility of Hyperpolarized ^{3}He MR Imaging[1]

David J. Niles, MS
Stanley J. Kruger, BS
Bernard J. Dardzinski, PhD
Amy Harman, MS
Nizar N. Jarjour, MD
Marcella Ruddy, MD
Scott K. Nagle, MD, PhD
Christopher J. François, MD
Sean B. Fain, PhD

Purpose: To quantitatively evaluate interday, interreader, and intersite agreement of readers of hyperpolarized helium 3 (HPHe) MR images in patients with exercise-induced bronchoconstriction.

Materials and Methods: This HIPAA-compliant, institutional review board approved study included 13 patients with exercise-induced bronchoconstriction. On two separate days, HPHe MR imaging of the lungs was performed at baseline, immediately after a 10-minute exercise challenge (postchallenge), and 45 minutes after exercise (recovery). Patients were imaged at two sites, six at site A and seven at site B. Images were analyzed independently by multiple readers at each site. Lung volume, ventilation defect volume, ventilated volume, and the number of defects were measured quantitatively, and the location of defects was evaluated qualitatively at site A. Interday and interreader agreement were evaluated by using the intraclass correlation coefficient (ICC), and intersite agreement was evaluated by using a modified Bland-Altman analysis.

Results: The ICC between days for ventilation defect volume, ventilated volume, and number of defects was at least 0.74 at both sites. The ICC for lung volume was greater at site B (0.83–0.86) than at site A (0.60–0.65). Defects seen in the same location in the lung on both days included 19.7% of those seen on baseline images and 29.2% and 18.6% of defects on postchallenge and recovery images, respectively. Interreader ICC for each measurement was at least 0.82 for each site. Analysis of intersite agreement showed biases of 612 mL for lung volume, −60.7 mL for ventilation defect volume, 2.91% for ventilated volume, and −6.56 for number of defects.

Conclusion: The reported measures of reproducibility of HPHe MR imaging may help in the design and interpretation of single- and multicenter studies of patients with exercise-induced bronchoconstriction.

[1] From the Department of Medical Physics (D.J.N., S.J.K., S.B.F.), School of Medicine and Public Health (N.N.J.), and Department of Radiology (S.K.N., C.J.F., S.B.F.), University of Wisconsin, 1111 Highland Ave, Room 1005, Madison, WI 53705-2275; and Merck Research Laboratories, West Point, Pa (B.J.D., A.H., M.R.). Received September 27, 2011; revision requested November 28; revision received March 15, 2012; accepted April 30; final version accepted July 11. Supported by and performed in collaboration with Merck (Whitehorse Station, NJ). **Address correspondence to** S.B.F. (e-mail: *sfain@wisc.edu*).

Exercise-induced bronchoconstriction (EIB) is the obstruction of lung airways after strenuous exercise. It is often, but not always, associated with chronic asthma (1,2). The bronchoconstriction is regionally heterogeneous and causes the development of spatially variable areas of low ventilation (ventilation defects) distal to affected airways (3). Symptoms of EIB are transient and self-limiting, reaching a maximum during the first 10 minutes after exercise and dissipating in 30–60 minutes (4,5). Testing for EIB typically involves spirometric measurements before and after a standard exercise protocol, and a decrease in forced expiratory volume in 1 second (FEV_1) of 10% or more is considered to be diagnostic of EIB (6).

Spirometric testing is a simple, low-cost method for diagnosis of EIB. However, spirometry results reflect FEV_1 for the whole lung and do not provide information on regional bronchoconstriction or ventilation. The size and location of defects may have important implications in determining disease severity or the efficacy of drug treatment. Several lung imaging techniques to measure regional ventilation are available, including high-resolution computed tomography (CT) (7), nitrogen 13 positron emission tomography (PET) (8), and hyperpolarized helium 3 (HPHe) magnetic resonance (MR) imaging. CT and PET are not ideal for longitudinal studies that involve multiple visits or imaging before and after exercise, particularly in younger patients, because they involve the use ionizing radiation. HPHe MR imaging involves the use of a biologically inert nonradioactive contrast agent that relaxes from the hyperpolarized state slowly relative to the length of the imaging experiment (T1 = 20–30 sec at 1.5 T). Therefore, HPHe MR imaging is useful for longitudinal imaging of lung ventilation in disorders such as EIB.

Although HPHe MR imaging is a promising approach for assessing lung function, the reproducibility of semiquantitative and quantitative HPHe MR imaging biomarkers has not been thoroughly evaluated. Images must be analyzed for metrics such as defect number, location, size, and severity, and this analysis is often based on the subjective judgment of a human reader, who can introduce an unknown degree of bias and variability. In addition, different research groups may use different methods or criteria for identifying defects. Three aspects of reproducibility that must be evaluated include *(a)* the agreement of a single reader's evaluations of images for ventilated volume–associated measures and defect locations in the same patient imaged on separate days; *(b)* the agreement among multiple independent readers who used the same visualization and measurement tools; and *(c)* the agreement among multiple independent readers who used different visualization and measurement tools at different imaging centers. This study is a quantitative evaluation of interday, interreader, and intersite agreement of HPHe MR imaging in patients with EIB.

Advances in Knowledge

- The intraclass correlation coefficients for hyperpolarized helium 3 (^{3}He) images collected on two separate days were 0.60–0.86 for lung volume, 0.77–0.89 for ventilation defects, 0.74–0.89 for ventilated volume, and 0.77–0.82 for number of defects, indicating high agreement among readers between days.
- The intraclass correlation coefficients for multiple readers of hyperpolarized ^{3}He images at a single site were 0.82–0.91 for lung volume; 0.91–0.95 for ventilation defect, 0.92–0.96 for ventilated volume, and 0.91 for number of defects, indicating high agreement among readers.
- Ventilation defects that were seen in the same location in the lung on images from both days included 19.7% of defects on baseline, 29.2% on postchallenge, and 18.6% on recovery images, respectively.
- The average bias for ventilated volume was 2.91% for readers of hyperpolarized ^{3}He images at the two imaging sites.

Implications for Patient Care

- The ventilated volume measurement derived from hyperpolarized ^{3}He MR imaging can help quantify reproducible changes in regional lung ventilation before and after exercise challenge, as reflected by an intraclass correlation coefficient of 0.74–0.89 for ventilated volume measurements on multiple days.
- The ventilated volume measurement can potentially serve as an imaging marker to help diagnose pulmonary disorders, make decisions about therapy intervention, and determine longitudinal progression of obstructive lung disease.

Materials and Methods

Patients

We received support for this study from Merck and equipment from GE Healthcare. Authors who had no affiliation with these companies had full control over the

Published online before print
10.1148/radiol.12111973 **Content code:** VA

Radiology 2013; 266:618–625

Abbreviations:
EIB = exercise-induced bronchoconstriction
FEV_1 = forced expiratory volume in 1 second
HPHe = hyperpolarized ^{3}He
ICC = intraclass correlation coefficient
N_D = number of defects
V_L = lung volume
V_D = ventilation defect volume
V_V = ventilated volume

Author contributions:
Guarantors of integrity of entire study, D.J.N., B.J.D., S.B.F.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, D.J.N., S.J.K., N.N.J., S.B.F.; clinical studies, B.J.D., A.H., N.N.J., M.R., S.K.N., S.B.F.; experimental studies, M.R., C.J.F., S.B.F.; statistical analysis, D.J.N., S.J.K., S.B.F.; and manuscript editing, all authors

Conflicts of interest are listed at the end of this article.

study and the data at all times. Merck required a 30-day review of the manuscript in advance of publication, but the scientific integrity and results were entirely under the purview of the principal investigators (S.B.F., N.N.J.). The study was conducted in accordance with Health Insurance Portability and Accountability Act regulations and was approved by the internal human patients review boards of both institutions. Thirteen adult patients with mild asthma and a history of EIB were included in this study (mean age ± standard deviation for the entire group, 24.4 years ± 7.0; men, 24.8 years ± 2.6; women, 24.2 years ± 8.4). All patients had received spirometric test results showing a decline of at least 15% in FEV_1 after performance of a 10-minute exercise challenge. The purpose of this study was to establish the robustness of HPHe MR imaging results as biomarkers of pulmonary function in patients with obstructive lung disease at two different imaging sites with different equipment platforms. Six patients were imaged at the University of Wisconsin (site A), and seven patients were imaged at the Robarts Research Institute, University of Western Ontario (site B).

MR Imaging before and after Exercise Challenge

Patients performed a 10-minute treadmill exercise challenge during two visits separated by 7–14 days. HPHe MR images were acquired 3.5 hours before exercise (baseline), immediately after exercise (postchallenge) and 45 minutes after exercise (recovery) (Fig 1). Spirometric measurement of FEV_1 and forced vital capacity was performed within minutes of each MR imaging examination to verify changes in pulmonary function. To control for diurnal variation in lung function, patients were imaged at approximately the same time of day for both visits.

Imaging was performed at site A by using a 1.5-T Signa HDx MR imaging system (GE Healthcare, Milwaukee, Wis) with a flexible-wrap single-channel volume coil (IGC Medical Advances, Milwaukee, Wis). At site B, a 3-T Signa HDx MR imaging system (GE Healthcare) was used with a rigid-body single-channel volume coil (Rapid Biomedical, Rimpar, Germany). Helium 3 (^{3}He) was hyperpolarized by using a Heli-Spin prototype commercial polarizer (GE Healthcare Biosciences, Durham, NC) and the spin-exchange optical pumping method described by Möller et al (9).

Figure 1

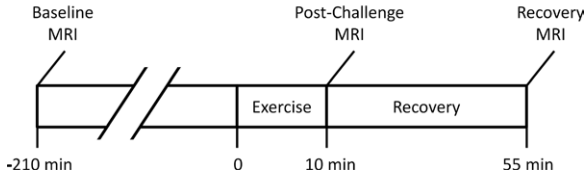

Figure 1: Timeline of exercise challenge and HPHe MR imaging examinations during each visit.

HPHe was prepared as a 4.5 mmol per liter dose mixed with nitrogen to a total volume of 1L in a polyvinyl fluoride bag (Jensen Inert Products, Coral Springs, Fla). The bag was purged of oxygen to slow T1 relaxation. While patients were in the imager, they inhaled the gas from functional residual capacity through a small plastic tube attached to the bag and were imaged during a 16–20–second end-inspiration breath hold. Blood oxygen saturation was monitored by means of pulse oximetry, and 100% oxygen was administered through a nasal cannula between imaging sequences to promote respiratory recovery. No adverse events were observed from gas inhalation or imaging procedures.

HPHe MR imaging at both sites used a fast two-dimensional gradient-echo sequence with the following parameters: repetition time, 3.6 msec; echo time, 1.1 msec; flip angle, seven degrees; number of sections, 12–16; field of view, 40 cm; acquisition matrix, 128 × 128; section thickness, 15 mm; and centric encoding. The acquisition matrix was reconstructed to a 256 square matrix, yielding a voxel size of 1.56 × 1.56 × 15 mm. At site B, proton images were acquired after patients inhaled a 1-L bag of ^{4}He from functional residual capacity with a three-dimensional spoiled gradient-echo sequence with a repetition time of 3.0; echo time, 1.1 msec; flip angle, 10 degrees; field of view, 40 cm; acquisition matrix, 256 × 224 (80% phase field of view); and section thickness, 10 mm.

Image Analysis

Images from all 13 patients were analyzed by two readers at site A (S.B.F. and D.J.N., with 10 and 2 years of experience, respectively) and three readers at site B (A.W., with 2 years and H.A. and S.C., both with 1 year of experience). Readers used the imaging approach of their respective sites and customized visualization and measurement tools developed in MATLAB (Mathworks, Natick, Mass). At site A, total lung volume (V_L) and the total volume of defects (V_D) were measured by means of manual segmentation of lung and defect boundaries of HPHe images (Fig 2). The number of defects (N_D) was automatically calculated in MATLAB on the basis of the connectivity of the segmented regions of defect. Defective regions that were not connected horizontally, vertically, or diagonally (eight connectivity) were counted as separate defects. At site B, defects were identified by consensus of the three readers, after which each reader independently measured V_L by manually segmenting registered proton images (Fig 2a) taken at baseline and V_D by manually segmenting HPHe images taken at each time point. Ventilated volume (V_V), a common imaging biomarker of ventilation defects (10), was calculated at both sites as $V_V = 100\% \times (V_L - V_D)/V_L$. Readers performed all image analysis while blinded to the randomized patient identification, day of imaging, and time point. With the exception of measurement of N_D at site B, each reader's measurements were independent.

Interday Agreement

Interday agreement refers to the similarity of measurements obtained at HPHe MR imaging examinations in the same patient on separate days and

Figure 2

a. b. c.

Figure 2: Coronal images of lungs show examples of segmentation steps performed at each site. **(a)** Segmentation of V_L from proton image acquired in patient at site B. **(b)** Segmentation of V_L from HPHe image of patient at site A. Ventilation defects appear as dark regions. **(c)** Segmentation of V_D from HPHe image.

evaluated by a single reader, or a single set of readers in the case of analysis by consensus. We evaluated the agreement of the location of defects and the values of V_L, V_D, V_V, and N_D obtained from image analysis. Two radiologists at site A (S.K.N. and C.J.N., with 7 and 9 years of experience, respectively) identified the location of individual defects by consensus and determined whether each defect appeared on images from both days. The reproducibility of defect location was then calculated for each patient and time point as the number of defects seen on images from both days divided by the total number of defects seen on images from either day. Images were excluded if no defects were seen on those of either day at a given time point (four of 78 examinations). Agreement of V_L, V_D, V_V, and N_D between the two days was evaluated separately for each of the five readers by using the intraclass correlation coefficient (ICC) and repeatability, with the latter defined as 2.77 times the standard deviation of the results of between the two imaging days in each patient (11). All ICC calculations in this study were tested for absolute agreement rather than for consistency.

Interreader Agreement

Interreader agreement refers to the similarity of measurements among all readers at a single site. Interreader V_L, V_D, V_V, and N_D at a single site were evaluated by using the ICC.

Intersite Agreement

Intersite agreement on V_L, V_D, V_V, and N_D values was evaluated by using a modified Bland-Altman analysis with the measurements by each reader treated as repeated measurements. An in-depth description of this method has been previously published (12,13). The presence of repeated measurements results in increased variance of measurement differences and a wider confidence interval (CI) than the traditional Bland-Altman analysis. If measured values at site A and site B are denoted as X and Y, respectively, the total variance of each method can be described as

$$\mathrm{Var}(X) = \sigma_t^2 + \sigma_{xI}^2 + \sigma_{xw}^2$$

$$\mathrm{Var}(Y) = \sigma_t^2 + \sigma_{yI}^2 + \sigma_{yw}^2$$

where σ_t^2 is the variance of the true values, σ_{xI}^2 and σ_{yI}^2 are the variances of each method by patient interaction, and σ_{xw}^2 and σ_{yw}^2 are variances of measurements in an individual patient for each method. The goal is to find the variance of the difference between single measurements by each method, Var $(X - Y)$, to determine the CI. For m_x and m_y repeated measurements at site A and B, respectively, this quantity becomes

$$\mathrm{Var}(X-Y) = \mathrm{Var}(\bar{X}-\bar{Y}) + \left(1-\frac{1}{m_x}\right)\sigma_{xw}^2 + \left(1-\frac{1}{m_y}\right)\sigma_{yw}^2$$

where $\bar{X}$ and $\bar{Y}$ are the average observed values among readers. The Bland-Altman bias, $\bar{D}$, is equal to $\bar{X}-\bar{Y}$ and the CI is calculated as $\bar{D} \pm 1.96 \times \mathrm{Var}(X-Y)$.

Results

Interday Agreement

Figure 3 shows typical images from examinations 7 days apart at site A (Fig 3, *A, B*) and site B (Fig 3, *C, D*). Defects appear with similar size and location in the lungs in each examination, as indicated by the arrows. On average, defects were seen at the same location on images from both days with different frequencies across the time points (Fig 4). Of the defects seen on baseline images, 19.7% ± 16.8% (range, 0%–40.7%) were present at the same location on the second day of imaging, compared with 29.2% ± 14.6 (range, 0%–56.5%) and 18.6% ± 16.9 (range, 0%–56.5%) of defects seen on post-challenge and recovery images, respectively. The frequency of repeated defects was significantly higher on

postchallenge images compared with those obtained at baseline and recovery ($P < .05$, Wilcoxon signed-rank test).

Interday ICC and repeatability values of V_L, V_D, V_V, and N_D for each of the five readers are shown in Table 1. The ICC values for V_D, V_V, and N_D were at least 0.74 in all readers and showed little difference between sites. However, the ICC of V_L showed a large difference between sites (0.60–0.65 at site A compared with 0.83–0.86 at site B), likely due to the use of different image sets for V_L measurement: HPHe images at site A and proton images at site B. As might be expected, interday repeatability varied inversely with interday ICC.

Interreader Agreement

Interreader ICC values for each site are shown in Table 2 and indicate good agreement for all independently assessed measures. Note that the N_D ICC values for site B were excluded because defects were identified by consensus at this site.

Intersite Agreement

Table 3 shows the average values for V_L, V_D, V_V, and N_D measured at each site for baseline, postchallenge, and recovery HPHe MR imaging examinations. V_L at site B was measured only at baseline and then replicated throughout examination time points. V_L (site A only), V_D, and N_D all increased after exercise, while V_V decreased, indicating symptoms of EIB. Although examination results from both sites show similar trends in measurements, there were biases between sites, as shown in the Bland-Altman analysis in Table 4. V_D and N_D were consistently higher at site B than at site A, while V_L and V_V were higher at site A. The measurement with the least bias was the V_V (−2.91%), which had an intersite CI of less than ±10.3%.

Figure 3

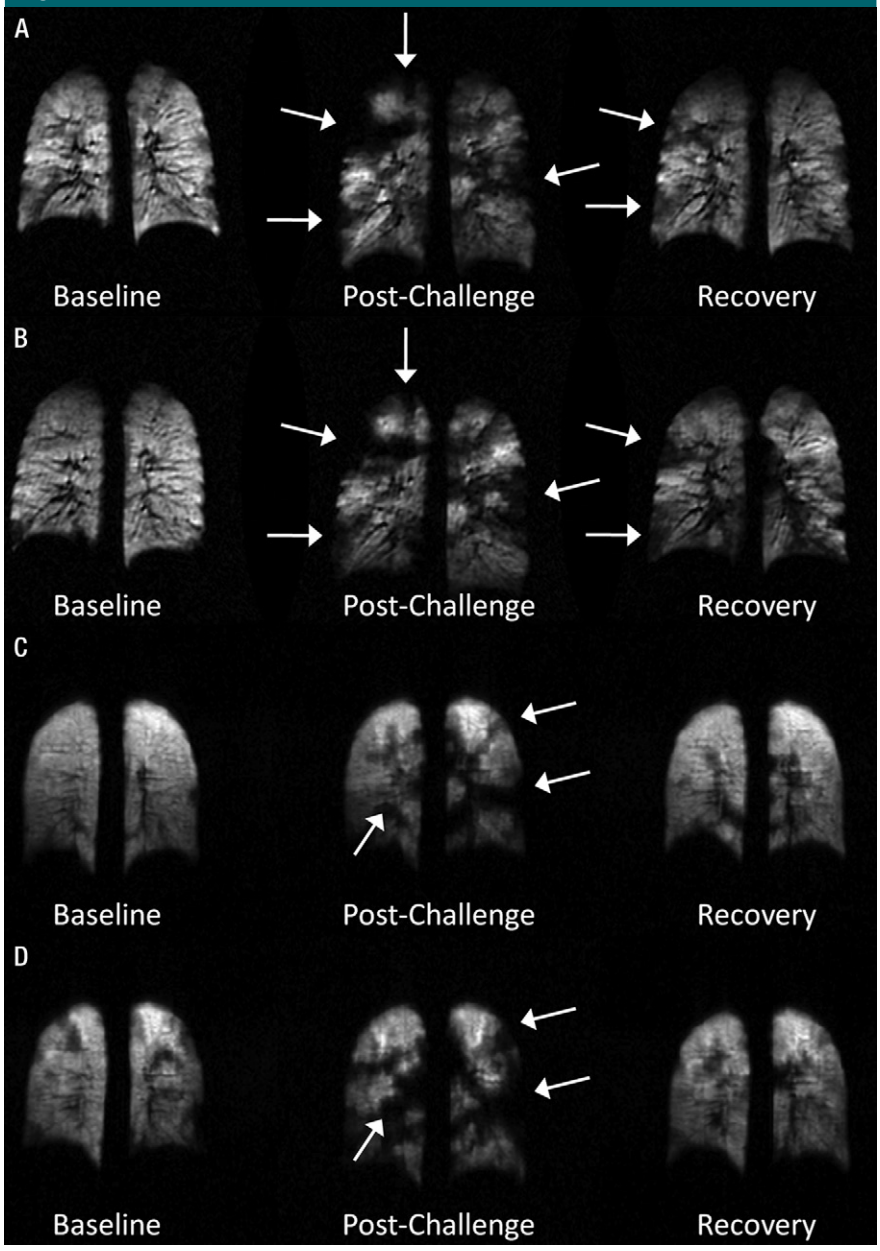

Figure 3: Typical coronal HPHe MR images acquired at baseline (left), postchallenge (middle), and recovery (right) time points for one patient on the, *A,* first and, *B,* second visits at site A and one patient on the, *C,* first and, *D,* second visits at site B. Defects (arrows) appear with similar size and location during each visit.

Discussion

Reliable estimates of reproducibility are necessary to compare HPHe MR imaging results between studies and to evaluate changes in ventilation after intervention such as drug treatment. Values for agreement and repeatability can be interpreted the minimum expected for the uncertainty of HPHe MR imaging measurements. To our knowledge, this study presents the first analysis of reproducibility of HPHe MR imaging of patients with EIB that compares results of image analysis at different sites.

The reproducibility of defects appearing at individual locations in the lungs throughout visits varied according to the extent of bronchoconstriction and/or the exercise state. Only 19.7% ± 16.8, and 18.6% ± 16.9 of

Table 1

Interday ICC and Repeatability for Each Reader

	Interday ICC					Interday Repeatability				
	Site A Readers		Site B Readers			Site A Readers		Site B Readers		
Measurement*	1	2	1	2	3	1	2	1	2	3
V_L (mL)	0.65	0.60	0.86	0.85	0.83	793	725	518	489	520
V_D (mL)	0.89	0.84	0.77	0.80	0.79	252	308	431	311	334
V_V (%)	0.89	0.88	0.74	0.75	0.75	5.37	5.92	10.5	8.53	9.01
N_D	0.77	0.82	0.78	0.78	0.78	17.3	16.1	21.6	21.6	21.6

Note.—Percentages for V_V were calculated as $100\% \times (V_L - V_D)/V_L$.

* (Units in parentheses apply only to repeatability columns. Other data are ICCs.

Table 2

Interreader ICC at Sites A and B

	Interreader ICC	
Measurement	A	B
V_L	0.91	0.82
V_D	0.91	0.95
V_V	0.92	0.96
N_D	0.91	. . .

Note.—N_D was not analyzed for site B because these defects were identified by consensus among readers.

Figure 4

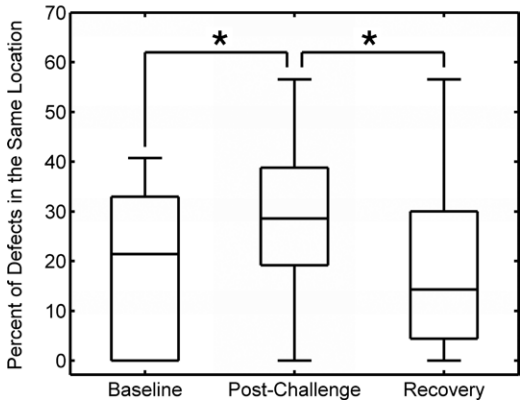

Figure 4: Box-and-whisker plots show percentage of defects observed in the same location in lung at both visits. Whiskers indicate largest and smallest values within 1.5 times interquartile range, and median values are indicated by horizontal lines in boxes. Significant differences (*) were observed between baseline and postchallenge values and between postchallenge and recovery values.

defects seen at baseline and recovery, respectively, appeared during both visits. This suggests that the location of defects that occur without exercise is highly variable. By comparison, 29.2% ± 14.6 of defects were visible in the same location at both visits after airway challenge with exercise. This observation is consistent with other work (14) showing recurrence of defects at the same locations in patients with asthma before and after methacholine challenge. Repeated bronchoconstriction in particular airways after exercise may reflect long-term remodeling, persistent injury, or sensitization of these airways.

The interday ICC of V_D, V_V, and N_D were all at least 0.74 for all readers, suggesting agreement among the five readers. The difference in the interday reproducibility of V_L between sites (ICC at site A, 0.60–0.65; at site B, 0.83–0.86; repeatability at site A, 725–793 mL; at site B, 489–520 mL) is likely attributed to the different source images readers at each site used to segment the lung boundary and is a limitation of the study design. Site B used proton images in which the lung boundary was clearly visible and not affected by ventilation. Site A, however, used HPHe images in which the lung boundary was less crisp and may have been completely obscured at the location of severe defects. This would require some degree of subjective estimation of the boundary location by the readers. The difference in methods between sites introduces variability in V_L, especially when defects occur at the base or the apex of the lung. Interestingly, the interreader ICC for V_L was high for both sites (Table 2). This suggests that readers at both sites arrived at similar estimates by using either type of source images. The repeatability of the V_V measurement at both sites is comparatively low in light of the higher repeatability values in V_L, perhaps because the V_D values are lower than those of V_L, which results in a relatively small dynamic range for V_V values compared with those of V_L.

Another revealing limitation of the analysis is that lung volume was noted to vary at different time points. Specifically, site B used proton images acquired only at the baseline time point to determine V_L and for normalization to calculate V_V at all subsequent time points. However, the values of V_L measured at site A (Table 3) indicated that patients have systematically different lung inflation volumes before and after the challenge ($P < .05$, between

Table 3

Mean Values for All Examinations for Each Site

	Baseline		Postchallenge		Recovery	
Measurement	A	B	A	B	A	B
V_L (L)	4.17 ± 0.57	3.73 ± 0.64	4.61 ± 0.74	3.73 ± 0.64	4.23 ± 0.58	3.73 ± 0.64
V_D (mL)	153 ± 191	210 ± 228	637 ± 433	668 ± 443	256 ± 326	344 ± 329
V_V (%)	96.4 ± 4.3	94.6 ± 5.1	86.3 ± 8.6	83.4 ± 10.3	94.0 ± 7.3	91.1 ± 8.1
N_D	15.7 ± 16.0	20.7 ± 18.2	32.1 ± 18.2	38.3 ± 22.0	20.8 ± 14.0	28.6 ± 18.5

Note.—Values are mean ± standard deviation. Percentages for V_V were calculated as $100\% \times (V_L - V_D)/V_L$.

Table 4

Bland-Altman Analysis Comparing Methods across Sites

Measurement	Mean Difference	Lower 95% CI	Upper 95% CI
V_L (mL)	612	178	1047
V_D (mL)	−60.7	−443	323
V_V (%)	2.91	−4.52	10.3
N_D	−6.56	−8.37	−4.75

Note.—Measurement differences were calculated as A minus B. Percentages for V_V were calculated as $100\% \times (V_L - V_D)/V_L$.

baseline and postchallenge), which is possibly due to obstruction of expiration and air trapping. This change in V_L would not be reflected by the single set of proton images captured at baseline in our study, which highlights the importance of acquiring proton images in conjunction with HPHe images at each examination, perhaps concurrently during a single breath hold (15), to determine V_L most accurately.

Perhaps not surprisingly, the Bland-Altman analysis revealed intersite differences in V_L, V_D, V_V, and N_D of varying magnitude depending on the measure. The largest difference was in the V_L measure, which was lower at site B by an average of 612 mL. The large V_L bias when we used proton images might be attributed to partial volume effects, particularly near the diaphragm, combined with perceived differences between positive and negative image contrast. The negative contrast of the lungs on proton images may cause the lungs to appear less inflated, and the positive contrast on HPHe images may cause an appearance of greater inflation, especially in areas of partial volume mixing. Thus, much of the bias in V_L likely depends on the inherent properties of imaging nuclei rather than site-specific effects such as the method of segmentation. Although the effects of positive and negative image contrast might be mitigated by inverting the grayscale of either proton or HPHe images, this bias may be important to note when comparing results from multiple studies and underscores the need for consistency in lung volume interpretation.

Another potential limitation of this study was the relatively narrow selection criteria, which focused on patients with both asthma and EIB. Reproducibility and robustness of the measures studied may not be as high in more heterogeneous disease processes in the general asthma population or in general obstructive lung disease. However, our results corroborate those of recently published studies (16,17) involving HPHe MR imaging of chronic obstructive pulmonary disease and cystic fibrosis. Mathew et al (16) reported an interday ICC of 0.74 for the V_V of HPHe MR imaging examinations performed 7 days apart. Kirby et al (17) reported an interreader ICC of 0.93 for V_V measured by two readers. In our results, both the intersite bias (2.91%) and CI (−4.52%, 10.3%) in V_V were relatively modest despite the differences in analysis methods, imager field strength, and coil hardware used. These recent results and our values suggest that V_V is a robust measure for multicenter studies of obstructive lung disease.

Finally, the image analysis methods used were semiautomated. Automated approaches such as k-means clustering (18) or statistical shape descriptor–based methods (19) would potentially mitigate the subjective elements of the analysis used in this study while providing substantial time savings. However, it is likely that a human observer would innately correct for image contrast and signal intensity differences between images from the two participating sites, such as those caused by different radiofrequency coils and, potentially, field strengths. Although authors of another study (20) have experimentally verified that the signal-to-noise ratio is independent of field strength in the 1.5-T and 3-T systems used in this study, differences in T2* effects were observed. A 3.3%–7.7% signal loss is expected at 3 T compared with that at 1.5 T at our acquisition parameters on the basis of published values of T2* (20,21). This small signal loss would most likely not affect a human reader, but it may be sufficiently large that a fully automated approach would require some level of intersite calibration for consistent results. Because images were both acquired and analyzed at two sites in this study, our

data set would be suitable for validating and investigating the robustness of an automated approach.

In conclusion, in this study we investigated the reproducibility of HPHe MR imaging in patients with EIB. Interday and interreader agreement by using a single method were generally high, but methodologic differences between the research groups produced differences in results for measures of V_L. Notably, the choice of image used to segment the lung boundary differed at the two sites and should be controlled for in future studies. Nonetheless, despite methodologic differences, V_V appeared to be a robust metric for evaluating ventilation in patients with EIB at both sites, and our results corroborated those obtained in studies of other obstructive lung diseases. The measurement of reproducibility in this study may guide the planning and interpretation of future HPHe MR imaging studies.

Acknowledgments: We acknowledge the financial support of Merck for this study. We thank Ryan Burton and Sandra Halko for conducting the exercise challenge; Kelli Hellenbrand, Jan Yakey, and Shayna McKay for study assistance; David McCormack, Grace Parraga, Giles Santyr, Andrew Wheatley, and Stephen Choy for their contributions to the analysis; and GE Healthcare for research support.

Disclosures of Conflicts of Interest: D.J.N. No relevant conflicts of interest to disclose. **B.J.D.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: Author is employed by and owns stock/stock options in Merck Sharp & Dohme. Other relationships: none to disclose. **A.H.** No relevant conflicts of interest to disclose. **N.J.J.** Financial activities related to the present article: Institution received grant from Merck. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **M.R.** Financial activities related to the present article: Institution received grant from Merck. Financial activities not related to the present article: Author is employed by Merck Research Laboratories. Other relationships: none to disclose. **S.K.N.** No relevant conflicts of interest to disclose. **C.J.F.** No relevant conflicts of interest to disclose. **S.B.F.** Financial activities related to the present article: Institution received grant from the National Institute of Health and the National Heart, Lung, and Blood Institute. Financial activities not related to the present article: Author is a consultant for GE Healthcare and Xemed. Author has grants/grants pending from GE Healthcare. Other relationships: none to disclose.

References

1. McFadden ER Jr, Gilbert IA. Exercise-induced asthma. N Engl J Med 1994;330(19):1362–1367.
2. Billen A, Dupont L. Exercise induced bronchoconstriction and sports. Postgrad Med J 2008;84(996):512–517.
3. Samee S, Altes T, Powers P, et al. Imaging the lungs in asthmatic patients by using hyperpolarized helium-3 magnetic resonance: assessment of response to methacholine and exercise challenge. J Allergy Clin Immunol 2003;111(6):1205–1211.
4. Brudno DS, Wagner JM, Rupp NT. Length of postexercise assessment in the determination of exercise-induced bronchospasm. Ann Allergy 1994;73(3):227–231.
5. Godfrey S, Silverman M, Anderson SD. The use of the treadmill for assessing exercise-induced asthma and the effect of varying the severity and duration of exercise. Pediatrics 1975;56(5 pt-2 suppl):893–898.
6. Weiler JM, Bonini S, Coifman R, et al. American Academy of Allergy, Asthma & Immunology Work Group report: exercise-induced asthma. J Allergy Clin Immunol 2007;119(6):1349–1358.
7. Arakawa H, Webb WR. Air trapping on expiratory high-resolution CT scans in the absence of inspiratory scan abnormalities: correlation with pulmonary function tests and differential diagnosis. AJR Am J Roentgenol 1998;170(5):1349–1353.
8. Harris RS, Winkler T, Tgavalekos N, et al. Regional pulmonary perfusion, inflation, and ventilation defects in bronchoconstricted patients with asthma. Am J Respir Crit Care Med 2006;174(3):245–253.
9. Möller HE, Chen XJ, Saam B, et al. MRI of the lungs using hyperpolarized noble gases. Magn Reson Med 2002;47(6):1029–1051.
10. Woodhouse N, Wild JM, Paley MN, et al. Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers. J Magn Reson Imaging 2005;21(4):365–369.
11. Bland JM, Altman DG. Measurement error. BMJ 1996;313(7059):744.
12. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8(2):135–160.
13. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17(4):571–582.
14. de Lange EE, Altes TA, Patrie JT, et al. The variability of regional airflow obstruction within the lungs of patients with asthma: assessment with hyperpolarized helium-3 magnetic resonance imaging. J Allergy Clin Immunol 2007;119(5):1072–1078.
15. Wild JM, Ajraoui S, Deppe MH, et al. Synchronous acquisition of hyperpolarised 3He and 1H MR images of the lungs - maximising mutual anatomical and functional information. NMR Biomed 2011;24(2):130–134.
16. Mathew L, Evans A, Ouriadov A, et al. Hyperpolarized 3He magnetic resonance imaging of chronic obstructive pulmonary disease: reproducibility at 3.0 tesla. Acad Radiol 2008;15(10):1298–1311.
17. Kirby M, Mathew L, Wheatley A, McCormack DG, Parraga G. Inter-observer reproducibility of longitudinal hyperpolarized helium-3 magnetic resonance imaging of chronic obstructive pulmonary disease [abstr]. In: Proceedings of the Eighteenth Meeting of the International Society for Magnetic Resonance in Medicine. Stockholm, Sweden; May 1–7, 2010: Abstract 2531.
18. Cooley B, Acton S, Salerna M, et al. Automated scoring of hyperpolarized helium-3 MR lung ventilation images: initial development and validation [abstr]. In: Proceedings of the Tenth Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, Calif: 2002.
19. Tustison NJ, Avants BB, Flors L, et al. Ventilation-based segmentation of the lungs using hyperpolarized (3)He MRI. J Magn Reson Imaging 2011;34(4):831–841.
20. Dominguez-Viqueira W, Ouriadov A, O'Halloran R, Fain SB, Santyr GE. Signal-to-noise ratio for hyperpolarized ^{3}He MR imaging of human lungs: a 1.5 T and 3 T comparison. Magn Reson Med 2011;66(5):1400–1404.
21. Deppe MH, Parra-Robles J, Ajraoui S, et al. Susceptibility effects in hyperpolarized (3) He lung MRI at 1.5T and 3T. J Magn Reson Imaging 2009;30(2):418–423.