Validating the distribution of specific ventilation in healthy humans measured using proton MR imaging

Rui Carlos Sá,1 Amran K. Asadi,1 Rebecca J. Theilmann,2 Susan R. Hopkins,1,2 G. Kim Prisk,1,2 and Chantal Darquenne1

1Pulmonary Imaging Laboratory, Department of Medicine, University of California, San Diego, La Jolla, California; and 2Pulmonary Imaging Laboratory, Department of Radiology, University of California, San Diego, La Jolla, California

Submitted 28 August 2013; accepted in final form 31 January 2014

Sá RC, Asadi AK, Theilmann RJ, Hopkins SR, Prisk GK, Darquenne C. Validating the distribution of specific ventilation in healthy humans measured using proton MR imaging. J Appl Physiol 116: 1048–1056, 2014. First published February 6, 2014; doi:10.1152/japplphysiol.00982.2013.—Specific ventilation imaging (SVI) uses proton MRI to quantitatively map the distribution of specific ventilation (SV) in the human lung, using inhaled oxygen as a contrast agent. To validate this recent technique, we compared the quantitative measures of heterogeneity of the SV distribution in a 15-mm sagittal slice of lung obtained in 10 healthy supine subjects, (age 37 ± 10 yr, forced expiratory volume in 1 s 97 ± 7% predicted) using SVI to those obtained in the whole lung from multiple-breath nitrogen washout (MBW). Using the analysis of Lewis et al. (Lewis SM, Evans JW, Jalalowayski AA. J App Physiol 44: 416–423, 1978), the most likely distribution of SV from the MBW data was computed and compared with the distribution of SV obtained from SVI, after normalizing for the difference in tidal volume. The average SV was 0.30 ± 0.10 MBW, compared with 0.36 ± 0.10 SVI (P = 0.01). The width of the distribution, a measure of the heterogeneity, obtained using both methods was comparable: 0.51 ± 0.06 and 0.47 ± 0.08 in MBW and SVI, respectively (P = 0.15). The MBW estimated width of the SV distribution was 0.05 (10.7%) higher than that estimated using SVI, and smaller than the intertest variability of the MBW estimation [inter-MBW (SD) for the width of the SV distribution was 0.08 (15.8)%]. To assess reliability, SVI was performed twice on 13 subjects showing small differences between measurements of SV heterogeneity (typical error 0.05, 12%). In conclusion, quantitative estimations of SV heterogeneity from SVI are reliable and similar to those obtained using MBW, with SVI providing spatial information that is absent in MBW.

respiration; specific ventilation; oxygen-enhanced MRI; multiple-breath washout; technique validation

THE SPATIAL DISTRIBUTION of ventilation is not uniform, even in healthy subjects, and is markedly heterogeneous in lung disease (24, 32, 34, 37). The time series of an inert gas [commonly nitrogen, helium, or sulfur hexafluoride (29)] washout of a multiple-breath washout (MBW) test can be used to determine the most likely distribution of specific ventilation (SV) within the entire lung (24). MBW consists of recording the breath-by-breath time series of the mixed expired nitrogen concentration while inhaling 100% oxygen; the washout data (mixed expired nitrogen concentration, as a function of expired volume) is then used to determine the most likely SV distribution, under some simplifying assumptions. This approach yields an estimate of the overall distribution of SV and a measure of heterogeneity, although no spatial information can be inferred.

Oxygen has been used as a contrast agent in proton magnetic resonance imaging (MRI) to assess regional ventilation in the human lung (13, 25), mostly in qualitative applications. Our laboratory has previously described a proton MRI technique to quantify regional SV [the local tidal volume to functional residual capacity (FRC) ratio] in a slice of the human lung (31). We have shown that this technique, termed SV imaging, or SVI, can reproduce the known vertical gradient in SV present in the human lung, in accordance with previous studies (23, 26, 31), but the technique has not been previously validated.

The SVI technique is similar to MBW in many ways: the approach is based on alternating breathing air and oxygen for blocks of 20 breaths and tracking signal changes during oxygen washin and washout. Oxygen (O2) is weakly paramagnetic and, when in solution in lung tissues, shortens the longitudinal relaxation time (T1) of the tissue [at 1.5 T, T1 ~ 1.3 s while breathing air, ~1.1 s breathing oxygen (5)]. The change in tissue O2 concentration, determining the local T1 change and thus the change in signal intensity in an appropriately timed inversion recovery data acquisition, is determined by local alveolar oxygen partial pressure (PAO2). The rate of change of PAO2 is in turn, determined by SV of that particular region of the lung during each oxygen washin and subsequent washout.

In its current implementation, a T1-weighted image of a single lung slice is acquired after each breath. Tracking signal intensity in each voxel thus provides a time series of regional washin/out. SV can thus be quantitatively determined by measuring the regional rate of change of the MRI signal intensity, following changes in inspired fractional oxygen content. By doing this on a voxel-by-voxel basis, SVI produces maps of SV in the lung, with an in-plane spatial resolution of the order of a few millimeters, for 10- to 15-mm-thick lung slices.

Since the heterogeneity of ventilation increases in healthy aging (37) and is markedly increased in lung disease (12, 24, 35, 36), it is highly desirable to have a reliable and valid technique for measuring the spatial distribution of ventilation. Different imaging modalities [PET (18), computed tomography (6), hyperpolarized MRI (22)] are capable of mapping the distribution of either SV or fractional ventilation in the human lung. None of these techniques, to the best of our knowledge, has been validated against nonspatial methods, such as MBW.

As a first step toward validation, the aim of this work was to establish the reliability and the validity of SVI in measuring the heterogeneity of SV in the healthy human lung. In the present work, we chose to study healthy subjects, because the assumption that a single lung slice is representative of the entire lung volume is likely valid in this population. To do this, we
compared the heterogeneity of SV measured by SVI to that measured by MBW.

**METHODS**

**Subjects**

This study was approved by the University of California, San Diego’s Human Subjects Research Protection program. Subjects participated after giving written, informed consent.

**Validation of SVI.** We studied 10 healthy subjects (4 women, 6 men), all between 22 and 52 yr of age. All subjects were recruited from the pool of subjects having previously undergone SVI (20, 31, 33) and were asked to return to perform MBW testing. The average interval between SVI and MBW tests was 21 mo (range 10–44 mo, 33) and were asked to return to perform MBW testing. The average from the pool of subjects having previously undergone SVI (20, 31, 33) and were asked to return to perform MBW testing. The average interval between SVI and MBW tests was 21 mo (range 10–44 mo, 33) and were asked to return to perform MBW testing.

**Subjects 1–10** participated in the validation portion of the study, subjects A–M in the reliability portion of the study. Bold indicate subjects who participated in both the specific ventilation imaging (SVI)-multiple-breath washout (MBW) comparison and SVI reliability. FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity.

**Table 1. Subject characteristics and pulmonary function data for all 20 subjects studied**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Age, yr</th>
<th>Height, m</th>
<th>Weight, kg</th>
<th>FEV₁, liters (%predicted)</th>
<th>FVC, liters (%predicted)</th>
<th>FEV₁/FVC, (%predicted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>44</td>
<td>1.65</td>
<td>57</td>
<td>3.16 (104)</td>
<td>4.06 (108)</td>
<td>0.78 (96)</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>52</td>
<td>1.86</td>
<td>113</td>
<td>4.50 (105)</td>
<td>5.36 (96)</td>
<td>0.84 (109)</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>27</td>
<td>1.78</td>
<td>73</td>
<td>3.96 (87)</td>
<td>4.92 (89)</td>
<td>0.81 (98)</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>40</td>
<td>1.75</td>
<td>93</td>
<td>3.49 (91)</td>
<td>4.28 (87)</td>
<td>0.82 (105)</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>42</td>
<td>1.58</td>
<td>71</td>
<td>2.62 (93)</td>
<td>3.45 (100)</td>
<td>0.76 (93)</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>22</td>
<td>1.75</td>
<td>68</td>
<td>4.33 (96)</td>
<td>5.71 (106)</td>
<td>0.76 (91)</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>29</td>
<td>1.73</td>
<td>62</td>
<td>3.63 (101)</td>
<td>4.36 (101)</td>
<td>0.83 (98)</td>
</tr>
<tr>
<td>8/A</td>
<td>F</td>
<td>39</td>
<td>1.63</td>
<td>66</td>
<td>3.08 (102)</td>
<td>3.65 (99)</td>
<td>0.84 (102)</td>
</tr>
<tr>
<td>9/B</td>
<td>M</td>
<td>34</td>
<td>1.78</td>
<td>83</td>
<td>3.84 (89)</td>
<td>4.98 (93)</td>
<td>0.77 (95)</td>
</tr>
<tr>
<td>10/C</td>
<td>M</td>
<td>30</td>
<td>1.70</td>
<td>62</td>
<td>4.42 (107)</td>
<td>5.27 (105)</td>
<td>0.84 (102)</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>24</td>
<td>1.76</td>
<td>92</td>
<td>3.71 (99)</td>
<td>4.34 (98)</td>
<td>0.85 (100)</td>
</tr>
<tr>
<td>E</td>
<td>M</td>
<td>26</td>
<td>1.86</td>
<td>93</td>
<td>4.57 (92)</td>
<td>5.28 (87)</td>
<td>0.87 (105)</td>
</tr>
<tr>
<td>F</td>
<td>M</td>
<td>26</td>
<td>1.73</td>
<td>68</td>
<td>4.36 (121)</td>
<td>5.06 (119)</td>
<td>0.86 (101)</td>
</tr>
<tr>
<td>G</td>
<td>M</td>
<td>33</td>
<td>1.80</td>
<td>81</td>
<td>4.03 (89)</td>
<td>5.26 (94)</td>
<td>0.77 (94)</td>
</tr>
<tr>
<td>H</td>
<td>M</td>
<td>22</td>
<td>1.79</td>
<td>84</td>
<td>5.17 (110)</td>
<td>6.55 (116)</td>
<td>0.79 (91)</td>
</tr>
<tr>
<td>I</td>
<td>M</td>
<td>28</td>
<td>1.88</td>
<td>82</td>
<td>4.79 (98)</td>
<td>5.17 (102)</td>
<td>0.78 (95)</td>
</tr>
<tr>
<td>J</td>
<td>M</td>
<td>24</td>
<td>1.86</td>
<td>98</td>
<td>5.03 (100)</td>
<td>5.79 (95)</td>
<td>0.87 (106)</td>
</tr>
<tr>
<td>K</td>
<td>F</td>
<td>24</td>
<td>1.60</td>
<td>61.8</td>
<td>3.09 (100)</td>
<td>4.16 (111)</td>
<td>0.74 (85)</td>
</tr>
<tr>
<td>L</td>
<td>M</td>
<td>21</td>
<td>1.77</td>
<td>76.8</td>
<td>4.74 (103)</td>
<td>6.23 (113)</td>
<td>0.76 (91)</td>
</tr>
<tr>
<td>M</td>
<td>F</td>
<td>22</td>
<td>1.66</td>
<td>57.5</td>
<td>3.99 (112)</td>
<td>4.46 (109)</td>
<td>0.89 (103)</td>
</tr>
</tbody>
</table>

**Data collection.** To avoid motion artifacts from cardiac movements, a single sagittal slice was selected within the right lung. The image slice location was selected in the midclavicular line, such that it presented the largest anterior-posterior dimension, while avoiding major hilar vessels. The slice position was referenced to the center of the subject’s spinal cord position to ensure that the same portion of lung was sampled in the second imaging session (reliability study). Two-dimensional T₁-weighted images were acquired in the supine posture, using a cardiac gated, global inversion recovery (inversion time, TI = 1,100 ms) single-shot fast-spin echo sequence, on a 1.5-T EXCITE HDi, TwinSpeed MRI system (General Electric Medical Systems, Milwaukee, WI). Images were collected using an eight-element torso coil, which provides higher sensitivity than the body coil built in to the scanner. Using the torso coil, signal-to-noise ratio for the lung was 6.6 ± 1.7 while breathing air, and 9.2 ± 1.8 while breathing oxygen. A half Fourier acquisition scheme (70-line acquisition, starting at −6 and extending to +63) was used to sample a field of view of 40 × 40 cm with a 256 × 128 resolution, resulting in an image acquisition time of ~300 ms. Images were reconstructed onto a 256 × 256 matrix, resulting in an in-plane resolution of ~1.6 × 1.6
mm, with a slice thickness of 15 mm (± 40-mm³ sampling volume). Successive images of the selected slice were acquired every 5 s, at FRC, with the subject voluntarily respiratory gating as described above. Images were acquired while breathing air for 20 breaths, and then 100% oxygen for the next 20 breaths. The air-oxygen cycle was repeated five times, and 20 additional 100% oxygen breaths were added at the end of the fifth cycle, resulting in a total of 220 images (~18 min). Expired tidal volume was measured simultaneously.

**SVI data analysis.** The time series of 220 images corresponding to the SVI sequence was imported into Matlab (Mathworks, Natick, MA), and quality control of the acquired raw images was performed. Image rejection was determined based on visual inspection of all acquired images, initially presented as a time-lapse series. Images with a visible displacement were selected and compared with a reference image by plotting a difference map; exclusion occurred when the lung border, typically the diaphragm, was found to differ compared with the reference. In the rare cases where the subject was not at FRC (average approximately nine rejected images per SVI run, 4.3%, with only one subject >10% rejected), images were removed from the series and replaced by an interpolated image, constructed from the preceding and following images. No image registration was applied to the data. A region of interest (ROI) selecting the lung field was manually drawn, and the subsequent analyses were restricted to the voxels inside the ROI. Data analysis was performed on a voxel-by-voxel basis; the time course of each voxel, MRI signal intensity vs. time, was used to estimate the local SV. All data analysis was performed using Matlab.

MBW, as described in further detail below, estimates the whole lung SV distribution, by determining a continuous, smooth distribution of SV arising from 50 parallel compartments that minimizes the error in predicting the decay in N₂ concentration during the washout. MBW considers 50 SV bins with center SV equally spaced in log₁₀, ranging from SV = 0.01 to 10. For the purpose of comparison with MBW, we introduced a modified SVI analysis, tailored to produce a SV map with identical binning as MBW.

To do this, the response of 50 independent SV compartments to the alternating sequence of inspired air and oxygen was simulated. Each simulated unit was attributed a SV, which, taking into account the alternating sequence of inspired air and oxygen was simulated. Each value of SV, was constructed. Since, in the measurement of SV described in Ref. 31, A matrix of the 50 simulated responses, one for each individual subject's tidal volume and the dilution normalization of tidal volume described below, resulted in a normal-

**Multiple Breath Washout**

**Experimental setup.** MBW was performed on a dedicated assembly similar to that used previously (12, 36). Briefly, the equipment uses a bag-in-box system with separate bags for inspired and expired gases, allowing for fast switching between inspired air or 100% oxygen. The inspiratory resistance was similar to the SVI setup, to minimize the difference in FRC between the two experiments. Flow and gas concentration were measured using a pneumotachograph (Fleisch no. 2), located in the wall of the bag-in-box system and a rapid-responding mass-spectrometer (Perkin-Elmer MGA 1100, Pomona, CA) with gas sampling at the mouth. Nitrogen, oxygen, and carbon dioxide concentration, as well as flow data, were collected at 200 Hz using an analog-to-digital converter and dedicated computer software. The system was calibrated for each subject. Calibration consisted of gas calibration with air and 100% oxygen, flow calibration using a calibrated 3-liter syringe (Hans-Rudolph) at a target flow rate of 0.5 l/s, and determination of mass-spectrometer transit time. The transit time was obtained by measuring the time required for a sharp puff of CO₂ to be detected by the mass spectrometer and was defined as the elapsed time between the time flow was detected and 50% of the rise in the CO₂ concentration. The transit time was ~600 ms, including both the lag time and the dynamic response time of the mass spectrometer. Ambient pressure and temperature remained stable during each subject's measurements and varied little from day to day (temperature = 24.5 ± 0.6°C, pressure = 760 ± 4 mmHg).

**Test maneuver.** The subject lay supine with the head tilted ~30° to the right, wearing a nose clip, while breathing through a mouthpiece attached to a filter (SDI diagnosis, or equivalent) and a non-rebreathing valve that was connected to the bag-in-box system described above. Using a metronome and a flowmeter, subjects were asked to breathe in a controlled and repeatable manner targeting a flow rate of 0.5 l/s and an inspiratory and expiratory time of 2 s each, resulting in a tidal volume of ~1 liter. Once the subject achieved a stable breathing pattern, the inspired gas was switched from air to 100% oxygen. The test lasted ~20–25 breaths, until the N₂ concentration fell below 2%. Tests were performed in triplicate, within ~30 min.

**Data analysis.** Data were analyzed using code implemented in Matlab (The Mathworks, Natick, MA) and Pascal. Mixed expired N₂ concentration was computed from flow and N₂ concentration data and used to determine the distribution of SV using the method described by Lewis et al. (24), as implemented by Prisk et al. (28). In short, this method assumes the lung is composed of 50 parallel compartments, each with a different SV value, uniformly spaced in log scale spanning the range from 0.01 to 10. As the distribution of SV uniquely determines the mixed expired N₂ concentration measured at the mouth, the most likely continuous SV distribution is determined by solving the inverse problem, using a weighted ridge regression algorithm (15, 24).

For each subject and repetition of the MBW test, the amplitude, center, and width of the SV distribution was computed, as for SVI, by fitting a log(Gaussian) function. The reported individual estimations are averaged from the MBW repetitions that pass the goodness of fit filter (R² > 0.75), to ensure that the log(Gaussian) function accurately describes the SV distribution.

**Comparing SVI with MBW: Volume Normalization**

MBW was performed in accordance to the European Respiratory Society/American Thoracic Society (ERS-ATS) guidelines (29), with a target flow rate of 0.5 l/s and a target volume of 1 liter. This is because interpretation of MBW results is complicated if performed outside the typical range of inspired volumes and controlled flows. SVI was performed as implemented in our laborat-
ory’s previous studies (20, 31), i.e., with the subject breathing at a paced rate (12 breaths/min), with tidal volume unconstrained (typically ~0.6 l/breath).

The different tidal volumes between the two techniques results in a difference in mean SV. To compare the two techniques, the individual distributions obtained using SVI were normalized to match the tidal volume recorded during the MBW session, an approach supported by previous work by Crawford et al. (8, 9), and comparable to the ERS-ATS suggested “volume compensation” for comparing pediatric subjects with different lung volumes (29). ERS-ATS guidelines suggest multiplying heterogeneity indexes (phase three slope, in the ERS-ATS suggestion) by the tidal volume to produce values comparable across a range of different lung volumes. Assuming a linear scaling of the SV distribution, the x-axis of the distribution obtained using SVI was multiplied by the ratio between the tidal volume measured from the MBW and SVI data. On a log10 scale, this corresponds to a parallel shift (right shift) of the SVI estimated distribution, changing the center but not the amplitude or width of the distribution, and results in identical width compartments for comparison purposes between the SVI and MBW.

Statistical Analysis

All data are presented as means ± SD. When data for all subjects are presented, SD refers to the intersubject variability (SD), unless otherwise specified. Except where otherwise indicated, paired t-test were performed to compare variables, with the null hypothesis (no effect) rejected when P < 0.05, two-tailed. All statistical analyses were performed using Prism (GraphPad, San Diego, CA).

**MBW-SVI comparison.** The amplitude, center, and width of the estimated SV histograms were compared using a paired t-test (paired by subject). A Bland-Altman comparison was also performed (3, 4). This analysis assumes no gold standard and allows determination of the average difference and limits of agreement between the two approaches. When available, the method was applied, taking into account the repeated measurements (3).

**Reliability.** Reliability was assessed using Pearson-product-moment correlation, the mean difference between the two measures, and the typical error (21). Bland-Altman analysis comparing the SVI repetitions, as well as the MBW repeated tests, was also performed. The intertest variability of SVI and MBW was computed as the standard deviation of each parameter over the retained repetitions.

**RESULTS**

**General Data**

Subject descriptive data and pulmonary function measurements for both the validation and the reliability study are presented in Table 1. All subjects studied had normal spirometry, indicating normal pulmonary function. The group studied in determining SVI reliability was, on average, younger than the subject population used for the validation study (27 ± 5 compared with 37 ± 10 yr old).

**Comparison between MBW and SVI**

Figure 1 presents a typical SVI map (Fig. 1A) and the SV distribution histogram for the imaged slice (Fig. 1B). Figure 2 presents a typical example tracing of N2 concentration and respiratory volume over time (Fig. 2A) and the distribution of SV obtained using MBW (Fig. 2B) for one of the three MBW repetitions, for the same subject depicted in Fig. 1. Figure 3 shows best fit log(Gaussian) distributions obtained from all SV distributions computed in one subject (same subject as in Figs. 1 and 2): three obtained using MBW, and two using SVI.

The average values for the normalized SVI distribution, averaged over all subjects, were as follows: amplitude 0.11 ± 0.02, center 0.36 ± 0.10, and width 0.46 ± 0.09, compared with MBW estimated amplitude of the distribution of 0.11 ± 0.01, center 0.30 ± 0.10, and width 0.51 ± 0.06. A paired t-test yielded no significant difference between the two estimations of the amplitude and width of the distribution (P > 0.9 and P = 0.15, respectively). The estimated center of the distribution was larger when estimated using SVI compared with MBW (P = 0.01).

A Bland-Altman analysis for the width of the distribution (heterogeneity) is presented in Fig. 4A. This analysis was used to determine the average difference between the two methods and to establish the limits of agreement between them. The average bias between MBW and SVI estimates of the heterogeneity of SV (Bland-Altman plot, Fig. 4A) was +0.05 (10.7%), with the individual differences ranging from −0.11

![Image](http://jap.physiology.org/)
to 0.21 (−20 to 36%) (Fig. 4A). The standard deviation of the individual bias was 0.09 (18.9%), compared with the average variability (SD of bias) between MBW repetitions of 0.08 (15.8%), range 0.11 to 0.13 (−20 to 26%) (Fig. 4B).

**Reliability of SVI Heterogeneity Estimation**

Results of the reliability testing showed a mean difference between the two SVI measures of −0.01 ± 0.01. The typical error (mean of the absolute difference between measurements) was 0.05, which corresponds to 12% of the value for SVI, consistent with many other physiological measurements (e.g., Refs. 7, 17). A Bland-Altman comparison of the two SVI experiments (Fig. 4C) resulted in a negligible bias < 0.01 (1.6%), and a SD of the bias of 0.06 (15.8%), similar to the SD of the bias observed for MBW. The Pearson product-moment correlation was moderate, $R = 0.74$, despite the low typical error, largely because of the homogeneous population studied. In comparison, the mean difference between MBW repetitions was 0.00 ± 0.01, and the typical error was 0.07 (corresponding to 13% of the MBW estimated SV heterogeneity).

**Quality Control**

On average, over all subjects, 2.3% of the voxels in the lung field did not pass the SVI statistical test of correlation with the air-oxygen driving function (3.1 and 1.4% in the validation and reliability groups, respectively). In all subjects but one, the range of excluded voxels was between 0.0 and 5%, and only one subject had >5% voxels rejected, due to poor signal-to-noise in that particular data set. This particular subject’s lung had a large antero-posterior dimension, resulting in poor signal-to-noise ratio (signal-to-noise ratio for the entire lung field was 3.5 while breathing air, compared with 7.0 on average for all other subjects), which was particularly low in the center of the lung field. Voxels that did not pass this filter were treated as missing data. In 2 out of 10 subjects, one MBW repetition failed to pass the $R^2$ filter for the goodness of fit, and those instances were excluded in the subsequent analysis of the subject.

**DISCUSSION**

This study showed that SVI produces reliable measurements of the heterogeneity of SV, and that those quantitative estimations are in close agreement with the estimation of SV heterogeneity obtained using MBW. The main focus of this paper was on the quantitative measure of SV heterogeneity produced by both methods. Given that changes in tidal volume and FRC change the center of the SV distribution, the heterogeneity of the distribution and not its center is the physiological parameter of interest in the SV distribution.

**Comparing MBW and SVI**

**Heterogeneity of SV.** The average difference between MBW and SVI estimations of the heterogeneity of SV (Bland-Altman plot, Fig. 4A) was small, 0.05 (10.7%), with a SD (bias) of 0.09 (18.9%). On average, the MBW SV heterogeneity estimation was higher than that obtained using SVI. The principle behind
MBW and SVI is the same, i.e., quantifying the temporal dynamics of a washin/washout, and both MBW and SVI follow similar exposures to inhaled oxygen: blocks of 20 breaths for SVI, ~20–25 breaths for MBW. Three main differences between the techniques are expected to cause a bias in the observed direction.

First, SVI, in its current form, is a single-slice technique, while MBW is a whole lung technique. Therefore, while the heterogeneity of SV estimated from MBW includes all heterogeneities present in the lung, the SVI estimation of SV distribution is based on a sample representing ~8% of the total lung volume at FRC. This may contribute to the higher SV heterogeneity in the MBW estimation that we observed.

Second, SVI, as currently implemented, requires 18 min of acquisition to acquire a single lung slice. The SVI estimation of regional SV corresponds to a temporal average of the local SV over the 18-min acquisition time. MBW on the other hand, is acquired over 2–3 min (20–40 breaths). The longer temporal averaging of the SV technique will tend to decrease the measured SV heterogeneity in SVI and may also contribute to the higher SV heterogeneity measured using MBW compared with SVI.

Third, the heterogeneity of the SV distribution measured using MBW and SVI is computed in different ways. SVI estimates the SV of individual voxels. In this way, if presented with a hypothetical noise-free, spatially uniform SV lung, the analysis of SVI would capture the uniformity of the lung. The mathematical analysis of MBW, as implemented by Lewis et al. (24), includes a smoothing parameter [used to stabilize the numerical solution of the inverse problem in the same way as in the multiple inert-gas elimination technique (15, 24)], which limits the resolution of the technique to quantify narrow peaks: SV distributions with heterogeneity below a certain width threshold are identified by the algorithm as a minimal detectable width (15, 28, 39). To determine the minimal detectable width imposed by the MBW analysis, we simulated nitrogen washouts, assuming homogenous lungs with different SVs in the range of interest (0.25–0.5), for 3-liter FRC and 150-ml dead space. For these simulated homogenous SV distributions (each a delta function), the MBW algorithm returned a SV heterogeneity of ~0.20 for physiological plausible SVs, as previously reported (28). Over all MBW tests of all subjects, the observed heterogeneity of SV was considerably higher than the minimal cutoff (average width 0.51, minimal MBW reported heterogeneity 0.35). Therefore, the SV heterogeneity observed in all subjects is in the range of applicability of the MBW analysis algorithm and represents physiological heterogeneity in SV within the lung.

Center and amplitude of the SV distribution. The center of the SV distribution was the only parameter found to systematically differ between the two measurement techniques; the center of the distribution as estimated using SVI was ~20% higher (P = 0.01) compared with MBW, following volume normalization. Several factors might contribute to this difference.

Fig. 4. Bland-Altman plots comparing the best estimate of SV heterogeneity for different methods/repetitions. The average of methods/repetition is plotted on the x-axis, and the difference between heterogeneity estimation of different methods/repetitions on the y-axis. In each subplot, the dashed line represents the average bias, and the dotted lines the bias ± SD (bias). A: MBW-SVI comparison, MBW produced SV heterogeneity estimations that are, on average, 0.05 (10.7%) above the corresponding SVI estimation. The bias standard deviation was 0.09 (dotted lines). B: comparison of SV heterogeneity between multiple repetitions of MBW. In this plot, the y-axis represents the difference between the third MBW repetition and the average of the prior two. Bias between MBW repetition was 0.01, and SD (bias) 0.08. C: comparison of SV heterogeneity between two SVI repetitions. Bias between SVI repetitions was 0.01, and SD (bias) 0.06.
First, the breathing maneuvers used in MBW and SVI resulted in different tidal volumes; therefore, measures of average SV were not directly comparable. For all subjects, average tidal volume during SVI was smaller than the breathing volume attained during the MBW maneuver. We corrected for the change in tidal volume between the two techniques, on a subject-by-subject basis, by normalizing by the ratio of the average tidal volume measured in both techniques. This SVI normalization approach is comparable to the “volume compensation” suggested by ERS-ATS guidelines for MBW (29) for comparing subjects with different lung volumes. In log x-axis coordinates, normalization resulted in a right shift of the SV estimated SV distribution (Fig. 1B). There is an implicit assumption in this normalization: for a given increase in tidal volume, different portions of the lung presenting different SVs will see their SV increase in the same proportion. Previous studies have supported this approach, reporting a linear increase in width of the SV distribution (heterogeneity) with increasing tidal volume (8, 9). The tidal volume correction (right shift) of the SV distribution measured using SVI depends on accurate tidal volume measurements between the two experimental setups. A systematic bias in tidal volume measurement could partially explain the observed difference between MBW and SVI. We do not believe this is the case, as care was taken to calibrate both instruments before each test using the same 3-liter syringe. The tidal volume normalization does not affect the amplitude of the SV distribution, which was found to be similar between MBW and SVI (0.5% difference, \( P > 0.9 \)).

Second, the in silico simulations of \( \text{N}_2 \) washout in homogeneous lung mentioned above also showed that the MBW algorithm underestimated the center of the SV distribution by \( \sim 11\% \) (for \( SV = 0.3 \)). This can partially justify the difference observed (\( \sim 20\% \)) between the MBW and SVI estimations of the center of SV distribution.

Third, the center of the SV distribution also depends on preinspiratory lung volume (8). All comparisons in this study were made with the subjects in the supine position, inspir ing from FRC, to maintain a repeatable preinspiratory volume. However, small changes in FRC might be present between the two techniques, as head position was different. The impact of small differences in FRC in the determination of the center of the SV distribution is expected to be small (8, 30).

Lastly, the two techniques sample different portions of the lung: the whole lung for MBW, and a volume representing \( \sim 8\% \) of FRC in the right lung. In the healthy normal lung, we expect a sagittal slice to be representative of the entire lung. However, the difference in sampled volumes can also be a factor contributing to the observed differences.

**Reliability of SVI and MBW in Estimating SV Heterogeneity**

The MBW and SVI techniques present similar reliability in estimating SV heterogeneity: SVI performed at two different times, showed a negligible average difference between the measurements of SV heterogeneity (i.e., width of the distribution) of 0.01 (1.6%) (Fig. 4C); SV heterogeneity computed from repeated MBW experiments presented an average bias of 0.01 (1.7%) (Fig. 4B). The variability between tests, both for MBW and SVI, is comparable, with a test-retest variability, SD (bias), of 0.07 (15.8%) for SVI and 0.08 (15.8%) for MBW. Possible sources of variability include both physiological and methodological contributions. Physiological sources of variability are common to both methods and include changes in the underlying SV distribution, breath-to-breath variability in inspired volume present in SVI and to a lesser extent in MBW, where flow and respiratory pace is imposed, and changes in FRC over time (time scales of days for SVI, minutes for MBW). Methodological sources of test-retest variability differ between the two methods, and include, for SVI, low signal-to-noise data, slight misregistration between consecutive images, and small spatial offset in slice selection between different tests. MBW is limited in its ability to detect narrow SV distributions (discussed above) and is sensitive to changes in breathing pattern and preinspiratory lung volume.

To assess the spatial reliability of the repeated SVI measurement, we performed an additional analysis, where the repeated SVI maps obtained in the reliability portion of the study were compared spatially. The analysis consisted of the following: 1) performing an affine transform between the two lung ROI, and applying it to the SVI maps, so they overlap spatially: 2) applying smoothing (as described in Ref. 20), with the single change that the local geometric mean was used as the scale of SV is logarithmic, to both SV maps to a spatial resolution of 1 cm\(^3\); and 3) computing the slope, intercept, and Pearson correlation of the linear relationship between the two SVI runs for each subject. The average Pearson correlation was 0.72 ± 0.15 (range 0.40–0.93). The slope averaged over all subject was 0.91 ± 0.34, and the intercept 0.05 ± 0.04. This amounts to a good spatial reliability of the SVI, especially when it encompasses any small differences between images from slice location.

The different time interval between MBW and SVI tests is a potential source of variability in estimating the validity and reliability of SVI. The comparison of SV heterogeneity measured using SVI and MBW, excluding the two subjects studied >40 mo apart, resulted in slightly higher average bias and similar SD (bias) as reported for all subjects. Similarly, the analysis of the reliability of SVI restricted to the subgroup of subjects studied within 24 h (\( n = 10 \)) produced similar average bias and SD (bias) compared with the entire group (including three subjects studied \( \sim 30 \) mo apart). The similar intratechnique variability for MBW and SVI, given the effort placed in minimizing the sources of error and variability, suggests that physiological variability is an important component of the overall intratest variability, and that SV distribution may change slightly over time.

**Quality Control and Limitations**

In the analysis of SVI, regions of low signal-to-noise ratio, regions of the lung with very low and very high SV, may result in erroneous estimation of SV. Due to the duration of the stimulus (20 oxygen breaths), regions of very low SV (\( SV < 0.01 \)) may not be quantifiable: over 20 oxygen breaths, low SV results in an increase in MR signal of the same order of magnitude as the noise in the regions of the lung farthest from the coil array. At high SV (\( SV > 1 \)), the sampling rate (1 image every 5 s) limits our accuracy to resolve regions that equilibrate very fast. This limitation is partly addressed by logarithmic separation between quantifiable SV bins. In addition, the quantification algorithm includes a built-in step of quality control.
that excludes regions where quantification has a large associated error. This step has the potential to misrepresent SV heterogeneity, by preferentially excluding the extremes of the distribution. We do not think this was the case in the present data, as the amount of rejected voxels was small (average 2.3%, and only in one subject did it exceed 5%).

Large pulmonary veins have the potential to transport to the imaging plane changes in O₂ concentration occurring in distant regions of the lung, thus potentially adding confounding information. We minimized this effect by selecting a midlung slice, lateral to the main pulmonary blood vessels, as determined from an initial anatomical set of localizer images, and deliberately avoided regions with large blood vessels.

Image registration was not performed in the data presented. Images where subjects failed to return to FRC were rejected and replaced by an average of neighboring images; this quality control step excludes evident misregistration between images. To estimate the impact of small misregistration errors between the acquired images, we have implemented the image registration scheme described in Refs. 2 and 20, based on registration of the lung border, on a subset of five subjects. Image registration resulted in minor (nonsignificant) changes in the amplitude (0.02), center (0.00), and width (0.04) of the distribution compared with nonregistered data. It is known that, in healthy subjects, regions with similar ventilation tend to cluster spatially (1); this might partially explain the resilience of the SV analysis to minor misregistration in the acquired images. In addition, SVI quantifies SV over five washin and washout cycles; this repetition provides resilience to isolated misregistration. However, spatial clustering might not be present in patient populations, and image registration may be a required data processing step for applying SVI to disease.

As mentioned above, single-slice SVI sampled ~8% of the lung. As gravity is responsible for most of the heterogeneity of SV in healthy subject (23, 26, 31), images were acquired in the sagittal plane in the right lung (maximizing the gravitational extent of lung sampled and moving away from the confounding effect of heart movements). In healthy subjects, a sagittal slice is thus representative of the entire lung. This is likely not the case in the presence of heterogeneous lung disease such as asthma (10, 11, 16, 19) or chronic obstructive pulmonary disease (14, 27). In the presence of heterogeneous lung disease, single-slice SVI may underestimate heterogeneity, as the sampled volume may not be representative of the whole lung. The results presented show that single-slice sampling accurately quantifies SV heterogeneity in healthy subjects. Generalization of SVI to heterogeneous lung disease will likely require the following: 1) whole lung SVI: the assumption of a single slice being representative sample of the whole lung will likely not hold in the presence of heterogeneous disease; 2) for applicability beyond clinical research, the acquisition time should be shortened; and 3) evaluation of the impact of potential large (>1 order of magnitude) regional changes in V˙A/Q occurring in the time frame of the SVI measurement. Such local changes are not present in the healthy lung, where the range of the V˙A/Q distribution is narrow (typical width ~0.3–0.4), and remains fairly unchanged when the subject is exposed to oxygen (38); if present in disease, large changes in V˙A/Q have the potential to introduce error in the SV quantification.

In conclusion, SVI allows for quantitative proton-MRI spatial mapping of SV in the healthy lung, producing repeatable estimates of the heterogeneity of the distribution, that are accurate compared with MBW and provide spatial information.

ACKNOWLEDGMENTS

This work benefited substantially from the technical contributions of Trevor Cooper (MBW software) and Janelle Fine (MBW hardware). The repeatability data were collected in different studies by the authors as well as Tatsuya Arai, Vincent Tedjasaputra, and A. Cortney Henderson.

GRANTS

This work was supported by National Heart, Lung, and Blood Institute (NHLBI) Grants R01 HL-080203, R01 HL-081171, and R01 HL-104118. Research work was supported by the National Space Biomedical Research Institute through National Aeronautics and Space Administration Grant NCC 9–58. A. K. Asadi was supported by NHLBI Grant F30 HL-110755.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES