Published in final edited form as:
Magn Reson Med. 2013 December ; 70(6): . doi:10.1002/mrm.24604.

A Variability Study of Regional Alveolar Oxygen Tension Measurement in Humans Using Hyperpolarized ^{3}He MRI

Hooman Hamedani[1], **Stephen J. Kadlecek**[1], **Masaru Ishii**[3], **Kiarash Emami**[1], **Nicholas N. Kuzma**[1], **Yi Xin**[1], **Milton Rossman**[2], and **Rahim R. Rizi**[1]

[1]Department of Radiology, University of Pennsylvania, Philadelphia, PA, United States

[2]Pulmonary, Allergy and Critical Care Division, University of Pennsylvania, Philadelphia, PA, United States

[3]University of Virginia, Department of Radiology, Charlottesville, VA, United States

Abstract

To presents the first systematic reproducibility measurement of alveolar partial pressure of oxygen (p_AO_2) in the human lung, regional variability is defined in terms of an intraclass correlation coefficient (ICC) between co-localized, same-subject measurements separated by one-week or couple of minutes (short-term). In addition, the repeatability of the average lung p_AO_2 is compared to that of the standard pulmonary function tests (PFT).

PFT and p_AO_2 imaging were performed on eight subjects: 4 nonsmokers (1 man, 3 women; 56 ± 1.7 years), 4 smokers (1 woman, 3 men; 52 ± 7.5 years) in three visits during two weeks. Regional variability was assessed based on a mixed-effects model and an ICC. The coefficient of variation (CV) of mean and standard deviation of p_AO_2 in three days was also compared to CV of PFT results.

Short-term regional reproducibility based on ICC was 0.71 and 0.63 for nonsmokers and smokers; respectively. The one-week variability was lower (ICC=0.59 and 0.47; respectively). The CV of whole-lung average p_AO_2 was significantly higher than that of FVC(forced vital capacity; P=0.02) but not from DLCO (diffusing capacity).

The smoker group shows more variability in p_AO_2 measurements both between experiments and in each individual p_AO_2 maps. p_AO_2 had a similar repeatability to DLCO.

INTRODUCTION

A variety of lung evaluation methods are in current use, which are sensitive to different aspects of pulmonary health. Among imaging techniqes, high-resolution chest-CT in particular can give excellent structural images, highlighting tumors and other disorders with a distinctive structural component. Imaging techniques to measure lung function, however, are not as well developed. CT-angiography is used clinically to detect blood clots and AVMs (arteriovenous malformation), and ventilation-perfusion scans can detect gross abnormalities of ventilation-perfusion mismatch. Other tests of lung function, such as pulmonary function tests, the six-minute walk test, cardiopulmonary exercise testing and methacholine challenge/bronchodilator response only measure overall lung function and do not give any insight into regional abnormalities.

Corresponding Author: Hooman Hamedani, University of Pennsylvania, Department of Radiology, 338 Stemmler Hall, 3450 Hamilton Walk, Philadelphia, PA, Tel: 215-662-6775, Fax: 215-746-8203, hooman@mail.med.upenn.edu.

A measurement of alveolar partial pressure of oxygen (p_AO_2) provides information about the quality of gas exchange in the functional units of the lung. Global p_AO_2 can be ascertained using two indirect techniques: helium washout and analysis of the mixed expired partial pressures of CO_2, O_2, and their arterial concentrations (1, 2). Yet these methods fail to provide enough information, as they lack the sensitivity to detect regional and small changes in the ventilation to perfusion ratio (VA/Q). Certain invasive techniques have been used in animals to overcome these regional limitations (3, 4). Such approaches, however, are not feasible in a clinical setting for human subjects.

Recently, hyperpolarized (HP) gas magnetic resonance imaging (MRI) has emerged to fill this void. This new technique, which is safe and sensitive, provides direct imaging of gas in the airways and parenchyma. The use of ^{3}He gas in particular has enabled imaging of both structural and functional aspects of the lung. Specifically, p_AO_2 imaging using hyperpolarized (HP) ^{3}He MRI has the potential to provide a quantitative and regional map of lung function (5–18), which can be a valuable tool for investigating the physiological changes in lung disease, the response to stress in the healthy lung, and innate differences among individuals (13, 19, 20). The p_AO_2 changes linked to disease state have been previously observed; with the aim of defining the significance of these changes and those of future investigations, this study presents a systematic short-term and repeat-visit variability measurement of p_AO_2 in the human lung. Regional variability is defined in terms of an intraclass correlation coefficient (ICC) between co-localized, same-subject measurements separated by one-week or several minutes (short-term test-retest). In addition, the repeatability of the average lung p_AO_2 is compared to that of the standard pulmonary function tests (PFT).

MATERALS AND METHODS

Subject Groups

All experiments proceeded under a protocol approved by the local Institutional Review Board with the subjects' informed consent and were Health Insurance Portability and Accountability compliant. Eight human subjects were divided into two independent cohorts: 1) four healthy nonsmokers who smoked less than 100 cigarettes in their lifetime, 2) four asymptomatic smokers with smoking history of at least 20 pack/years.

Pulmonary Function Tests

Spirometry, plethysmography and diffusing capacity for carbon monoxide testing was performed on all individuals just prior to each imaging session according to ATS/ERS 2005 guidelines (21) in a hospital pulmonary function laboratory.

Imaging Acquisition

A whole-body 1.5-T MRI scanner (MAGNETOM Sonata, Siemens Medical Solutions, Malvern, PA) using an 8-channel chest coil (Stark Contrast, Erlangen, Germany) tuned to the ^{3}He resonance frequency of 48.48 MHz was used for all ^{3}He imaging.

Regional p_AO_2 was measured from a set of successive HP ^{3}He spin-density signals, based on the linear relationship between the O_2-induced $1/T_1$ decay rate of ^{3}He spin polarization and the local oxygen concentration (22). A multi-slice, paired single-breath acquisition scheme involving an interleaved sequence was used for p_AO_2 imaging based on an earlier study (18). This multi-slice scheme utilizes four series of images for each slice; four images of the twelve coronal slices (A to L) were produced continuously in an "AABBCC…KKLLAABBCC…KKLL" pattern following the interleaved acquisition scheme. Each experiment utilized twelve coronal slices imaged using a gradient-echo imaging pulse-

sequence with the field of view (FOV) = 40 × 30 cm^2, slice thickness (ST) = 13 mm, inter-slice gap = 20% ST, $\alpha_{nominal} = 5°$, matrix size = 48 × 36 pixels, and $T_R/T_E = 6.69/3.1$ ms. This multislice p_AO_2 imaging took approximately 12 seconds with 8.3 × 8.3 × 15.6 mm^3 voxel-size resolution.

A commercial prototype polarizer (IGI 9600.He, GE Healthcare, Durham, NC) hyperpolarized the imaging gas ($^3He:N_2$ = 99.19:0.81, Linde, Branchburg, NJ) through spin-exchange collisions with optically pumped rubidium atoms. 15 hours of optical pumping produced polarization levels of 25–35%. In all experiments, hyperpolarized ^3He gas was diluted with medical-grade nitrogen gas (3:1). Dispensed into a Tedlar plastic bag (Jenset Inert Products, Coral Springs, FL), the gas mixture was brought to the MRI scanner for use.

Human Experiments

Testing on each subject occurred on three days over a two-week period. On each day of the experiment, each subject underwent two p_AO_2 measurements in the supine position performed 5 ± 3 minutes apart. The temporal ordering of slices in the two same-day acquisitions was reversed, with the first proceeding in the anterior-toposterior (A→P) direction and the second in the P→A direction. Throughout testing, subjects' vital signs were constantly monitored under the supervision of a physician. Images were acquired during a breath-hold following the subjects' inhalation of a mixture of HP ^3He, N_2, and O_2 at a prescribed ratio with F_iO_2 at approximately 21%. The inhaled gas volume was 12% of total lung capacity (TLC) as measured by plethysmography. The Tedlar bag and hyperpolarized gas mixture (9.5% TLC) was transferred to the bore of the MRI scanner and connected to a three-way pneumatic valve along with a bag containing O_2 gas (2.5% TLC). Subjects were instructed to breathe normally over three breath cycles with an inhale:exhale ratio of 3:4 at a uniform rate before inhalation. Then, a three-way valve was actuated and the subject inhaled the full contents of both bags simultaneously. Imaging began three seconds after the initiation of the breath-hold in order to ensure appropriate post-inspiratory mixing of gases and to minimize gas redistribution and flow artifacts in the measurements. The total breath-hold lasted ~15 seconds. The subject then exhaled forcefully back into the bag upon completion of the imaging. The exhalation also enabled the measurement of end-tidal O_2 and CO_2 using the respiratory gas analyzer.

Image Analysis

Custom software developed by the MATLAB environment (MathWorks, Inc., Natick, MA) was used for image analysis on a slice-by-slice basis. Using each voxel's signal-intensity drop in consecutive images, a least-square fit to a model of O_2-induced $1/T_1$ decay rate of ^3He spin was performed in order to calculate p_AO_2 for each voxel in all slices (18). Twelve p_AO_2-maps were generated for each subject. The mean (μp_AO_2) and standard deviation (σp_AO_2) of all voxels' p_AO_2 over the twelve slices were calculated for whole-lung comparisons.

For regional comparisons, an affine-based coregistration (23) was performed on p_AO_2 maps of each subject's six measurements in three days. Then, the voxels were binned into 12×9×5 isotropic (3.3×3.3×3.1 cm^3) regions of interest (ROIs) by summing 4×4 inplane voxels and 2 voxels in the slice direction. The most anterior and posterior slices were excluded from analysis due to relatively low signal and motion sensitivity.

Statistical Analysis

The coefficient of variation (CV) of the μp_AO_2 and σp_AO_2 was compared to the CV of the PFT results using a paired t-test to discern the relative repeatability of these measurements

(the CV of a quantity refers to the standard deviation of the 3 days' measurements divided by its average).

Regional variability analysis was performed using STATA 12 (StataCorp, College Station, TX). A mixed-effect regression model was fit to the p_AO_2 values of each corresponding ROI. Using a maximum likelihood technique during model construction, the models were fit with the backward elimination procedure based on likelihood ratio tests and the Akakie information criterion (24). Once the final fixed covariates were determined, the final model was reconstructed using the restricted maximum likelihood technique (25). Two models were fit according to the variance structure presented in the mixed-effect model below; the first was designed to estimate the weekly follow-up variability and second to evaluate the test-retest repeatability (~5 minute apart). The latter exploits the observation that the image acquisition direction represents a small effect and serves as a surrogate for measurements taken in short succession with opposing slice acquisition direction. We therefore derive the proportion of variability in measurements between the two acquisition directions to estimate the test-retest repeatability. For this short-term repeatability, the second day's back-to-back measurements (~ 5 minutes) were used. An Intraclass Correlation Coefficient (ICC) was derived from ratios of the calculated variances. Formally, the ICC represents the within-ROI correlation; conceptually it can be thought of as the proportion of total measurement variance that can be explained by subject identifier alone.

To compare the variability of short- and long-term measurements, the standard error of the difference between the weekly and minute intraclass correlation coefficients was estimated using the bootstrap sampling method. 500 samples were used for this calculation. We obtained the 95% confidence intervals for the difference using the normal approximation and the estimated standard errors. The differences were assumed to be statistically significant if the 95% confidence intervals did not include zero.

Mixed Effect Model

Tested fixed effects include a subject group factor, a gradient term by slice, image acquisition direction, weight, sex, and BMI. The derived variances were used to estimate the follow-up variability and the test-retest repeatability. It seems reasonable to assume that p_AO_2 values will be more correlated within subjects than between subjects, and that p_AO_2 values will be more correlated within slices than between slices (because slice data were obtained independently). We therefore account for random-effects due to subject, slice, and ROI in our regression model and assume that the ROI term is nested within the slice term, which is nested within the subject term.

For this study, all statistical tests used a significance level of $\alpha = 0.05$, and distributions were deemed significant if $P < 0.05$. Based on (26), an ICC<0.4 indicates *poor* repeatability, 0.4 – 0.80 is *fair to good* and over 0.80 is assumed *excellent* repeatability.

RESULTS

Human Experiments

Demographic characteristics are provided in Table 1 for the 8 subjects enrolled in the study (4 nonsmoker and 4 smoker). The average and standard deviations are also summarized in the table. Subjects' age, weight, body mass index (BMI) and total lung capacity (TLC) averages were not significantly different between the two groups.

During all the imaging studies, we monitored blood pressure, heart rate, and oxygen saturation and observed no adverse events. Oxygen saturation level did not fall below 90% in any of the subjects during the breath-hold. One subject (subject 7) in the smoker group

was not able to hold her breath and exhaled before the imaging on the first experiment of the first day. All the subjects tolerated the 15-second breath-hold, although one smoker (subject 5) was not able to remain motionless during the image sequence. Some subjects did not inhale the full contents of both oxygen and helium bags in their first attempt, but improved in the follow-up experiments. Three subjects (subject 2, subject 4 and subject 5) still had some difficulties either in their second or last day experiments, as was indicated by few gas remaining in the bags before imaging.

p_AO_2 maps

Figure 1 displays p_AO_2 maps and histograms for two representative subjects, a healthy non-smoker (subject 1) on the top and an asymptomtic smoker (subject 5) on the bottom. In this figure, each series of both p_AO_2 maps and ventilation images are listed. The most anterior and posterior slices are not shown for brevity. The histogram displays the p_AO_2 values of the whole-lung for each subject, which consists of all voxels from all slices. The p_AO_2 mean values and standard deviation (dispersion) are also listed for each slice as well as the whole-lung.

Whole-lung observations and repeatability

Whole-lung μp_AO_2 and σp_AO_2 (mean ± standard deviation) are listed in Table 1 for all of the A→P studies. The standard deviation is a heterogeneity marker for each subject. The highest μp_AO_2 (125.2 Torr) was observed in subject 4 (nonsmoker) and the lowest value (71.4 Torr) was observed in subject 6 (smoker). The lowest σp_AO_2 was 19.5 Torr in subject 4 (nonsmoker) and the highest σp_AO_2 was 42.9 Torr in subject 7 (smoker). Three days' average μp_AO_2 in smokers was lower than that of nonsmokers (Smokers: 87.5 ± 11.1 Torr, Nonsmokers: 103.0 ± 17.0) and three days' average σp_AO_2 in smokers was significantly higher than that of nonsmokers (Smokers: 36.2 ± 4.3 Torr, Nonsmokers: 27.7 ± 5.6 Torr, $t(6) = 2.39$; $P = 0.05$). The medians of the distribution (shown by small triangles in Figure 1) were in the range of 110–130 Torr for all representative subjects.

As shown in Figure 2, we assessed between three-days' repeatability of μp_AO_2 and σp_AO_2 for all subjects using a linear regression model. The slope of the regression for μp_AO_2 for all subjects and days was 0.94 and for σp_AO_2 the slope was 0.85. The corresponding test-retest Pearson correlation coefficients between day 1 and day 2 were significant for both μp_AO_2 ($r = 0.91$; 95% CI: 0.49, 0.99; $P < 0.01$) and σp_AO_2 ($r = 0.67$; 95% CI: −0.16, 0.95; $P = 0.05$). The correlations were stronger between day 2 and day 3; μp_AO_2 ($r = 0.95$; 95% CI: 0.72, 0.99; $P < 0.001$), σp_AO_2 ($r = 0.95$; 95% CI: 0.73, 0.99; $P < 0.001$).

Table 2 summarizes the results of the three PFT tests on each individual (mean ± standard deviation). The repeatability of the μp_AO_2 and σp_AO_2 was compared to the repeatability of the FVC, FEV_1, $FEF_{25–75\%}$, TLC and the DLCO based on the measurement's CVs and is displayed in bottom portion of Table 2. The t-test results between PFT results and global p_AO_2 values show that the repeatability of the p_AO_2 was significantly less than the FVC and FEV_1 but did not differ significantly from the repeatability of the DLCO and $FEF_{25–75\%}$.

Slice distributions

As can be observed in the representative nonsmoker (Figure 1), the anterior slices show higher p_AO_2 values. These values decrease monotonically from slice to slice. This monotonic gradient along the slices can be observed in the representative smoker subject as well, and was also seen in the subjects not depicted (see summary plots of Figure 5b).

Figure 3 shows the p_AO_2 distribution for slices in A→P direction for the two representative subjects. Superimposed histograms obtained on the three different days are also shown for

each subject. Each row of box-plots shows an individual and repeatable slice-dependent pattern and each subject displays a visually apparent consistency among overall histograms.

Figure 4 compares p_AO_2 maps of each slice from the other two days of testing in the nonsmoker and smoker representative subjects (Figure 1) and the histograms of the three days. Similar visual consistencies can be seen, particularly in the patterns of elevated and low p_AO_2.

Mixed-effect model

To formally test the empirical observations, data was fit using a mixed-effect linear model for follow-up and short-term test-retest studies. Spatially binned p_AO_2 data (isotropic ~32-cm^3 cubes as introduced in the Methods section) were used for this purpose. The summary box plots of Figure 5 shows binned p_AO_2 data versus imaging day, slice number and imaging direction. Regression and random-effect results are presented in Table 3 for weekly follow-up and in Table 4 for short-term test-retest measurements. All effects not listed were found to be insignificant. The constant fixed-effect coefficient is the average p_AO_2 of the binned data among all subjects in all experiments. Smokers had lower average pulmonary oxygen levels than nonsmokers (−33.57 and −23.38 Torr for long-term and short-term, respectively). There is a slice-dependent gradient with oxygen levels falling for each binned-slice from anterior to posterior (−10.56 and −12.83 Torr). The quadratic slice terms in both long-term (1.14 Torr) and short-term (1.58 Torr) cases, shows that the oxygen levels do not drop linearly but bottom off at the posterior slices. There is a cross term between slice and smoker (4.99 and 4.30 Torr, for long-term and short-term, respectively) which suggests that the gradient is less strong in smokers.

The nonsmoker and smoker random effect residuals in short-term studies are 12.95 and 15.68 Torr, respectively and the corresponding values for long-term studies are 14.60 and 18.84 Torr. Subject variability was small at 13.13 Torr for long-term and 14.24 Torr for short-term, and of interest was the observation that the variation between slices (3.57 and 4.83 Torr) was much smaller than the variation (11.12 and 13.65 Torr) between ROIs within a slice.

The subject, slice, voxel, and residual errors were used to estimate the conditional ICC(subject, slice, ROI), for the nonsmoker and smoker subjects, yielding 0.591 and 0.465, respectively for long-term measurements. These values represent the floor estimate of follow-up variability for an individual ROI oxygen measurement on the week time-scale.

For ~5-minute test-retest studies, the residual errors dropped, demonstrating that there is less variability in an individual ROI oxygen measurement over the minute time scale than over weeks. The improved ICC for the nonsmoker and smoker are 0.71 and 0.63, respectively for short-term repeatability; the significance of this improvement is confirmed by the fact that the 95% confidence intervals do not overlap. The bootstrapping results are summarized in Table 5.

DISCUSSION

The main finding of this study is a determination of the temporal repeatability among co-localized, same-subject p_AO_2 measurements and its relationship to the variability of the average p_AO_2 across subjects. This variability metric, summarized as the ICC, can be considered to show a *fair to good* (26) repeatability among healthy smokers and nonsmokers (ICC = 0.465 and 0.591, respectively) when comparing measurements separated by a week.

Several effects may limit the one-week repeatability. Chief among these is a true variation in the p_AO_2 at different times. In an effort to quantify this physiological variation, we compared one-week to short-term variability; the latter is likely to be minimally affected by changes in physiological state. We note that assuming minimal dependence on the order of slice acquisitions, we may take the A→P and P→A measurements (separated by approximately five minutes) as representative of short-term repeatability; these will show a higher ICC to the extent that physiological change determines variability in the p_AO_2 measurements. We found that the fixed-effect coefficients changed insignificantly, but the residuals dropped significantly, corresponding to a short-term ICC of 0.627 and 0.711 in smoker and nonsmoker. This demonstrates that there is less variability in individual ROIs over minutes than over weeks, and suggests that a significant amount of variability can be attributed to true physiology-dependent p_AO_2 changes between visits.

Furthermore, it is important to note that the between-visit variability is significantly larger in the smoker group than in the nonsmoker group. Because the experimental conditions are fixed by design, there is minimal chance that variation of other factors would change drastically between subject groups. This provides additional support for the hypothesis that temporal variability, particularly in the smoking subjects, is substantially physiological in origin.

Nonetheless, there are several other potential sources of measurement variability. As is discussed in (18), the measurement uncertainty is a strong function of signal-to-noise (SNR). Among the experiments in this study, the average SNR for ROIs were 425, corresponding to an expected random error of 2.5 Torr, or approximately 1/6 of the observed p_AO_2 distribution width in nonsmokers. It should be mentioned that the average SNR for individual voxels was much less than the abovementioned value and as a result the variability in the reported whole-lung dispersions that were based on all voxels, contributes more variability due to noise (~ 10–15 Torr).

Another potential source of measurement error is the diffusion or flow of gas between regions of interest over the course of the measurement. Flow into or out of the region of interest can mask or mimic the effect of oxygen with confounding signal dynamics. In healthy subjects, the majority of signal comes from gas contained within acinar regions, for which both diffusion and pressure-induced flow are expected to be very small. In accordance with (18), a 20% inter-slice gap was included to minimize any residual errors; nonetheless, non-negligible error may exist in airways, in severely emphysematous regions, and whenever air trapping causes pressure gradients to persist for many seconds into a breath hold.

In a similar manner, physical motion of the subject or the lung (e.g., near the beating heart) leads to coregistration uncertainty and erroneous measurements (17, 18). We ameliorated this effect in the analysis by choosing relatively large regions of interest, and believe that severe registration errors would have been detectable as a significantly larger variability in regions near the heart and diaphragm. This was not observed. Nonetheless, the effect of gas flow and motion on the measurements is not yet quantifiable.

Finally, there is potential for measurement errors related to the delivery of gas to the subjects and the subjects' ability to hold their breath in a reproducible manner. As opposed to some previous studies (12, 17), subjects breathed a normoxic gas mixture to avoid the potential for significant physiological response to hypoxia (3). However, to limit premature signal loss, the gases were kept in separate bags until administration and delivered together through a single mouthpiece. In a few cases, subjects stopped inhaling before both bags were completely empty, presumably in response to additional resistance encountered as the

bag is deflated. This apparent difficulty suggests the use of sufficient practice prior to the measurements and also criteria for exclusion of erroneous measurements, in a manner similar to the criteria used to evaluate the validity of FEV_1 measurements (21). Subjects were instructed to inhale and exhale normally three times right before inhaling the contents of the bags in order to begin inhalation from the true baseline functional residual capacity. Subjects who inadvertently exhaled more or less forcefully in preparation for hyperpolarized gas administration may skew the results. To avoid or identify these types of error, a real time inspiratory/expiratory measurement is desirable, which was not used in this study. Similarly, some subjects experienced difficulty controlling the breath-hold and involuntarily inhaled or exhaled slightly. For example, during the breath-hold, one subject (subject 5) was seen shaking while trying to hold his breath, which may explain unusually the large measurement variability between visits. A real-time monitoring scheme would justify rejection and repetition of these measurements as well.

In addition to introduction of ICC as a measure of regional variability, the one-week repeatability of the whole-lung p_AO_2 mean and dispersion were also assessed and are evaluated in the correlation plots of Figure 2. Both μp_AO_2 and σp_AO_2 show reasonable repeatability. These metrics provide a useful comparison to previous works and highlight a measure of reproducibility that is not limited by SNR or inter-ROI gas redistribution.

Beyond the potential to characterize subjects by their highly repeatable μp_AO_2, it is notable that the p_AO_2 histogram contains visually apparent features that are highly repeatable as well. In Figure 3, the p_AO_2 histogram shows clear differences between the two subjects; in addition to a distinct mean and dispersion, each subject is characterized by a distinct skewness and other qualitative features as well. As with the histogram width, skewness appears to be larger in the smoker cohort, although a larger study will be needed to demonstrate the significance of this effect.

In addition, we can observe a visual consistency in the box plots of Figure 3 for each subject, suggesting a unique and repeatable A→P or gravitational dependence of p_AO_2 for each subject. This trend is also present in subjects not included in Figure 3 (see Figure 5b). Interestingly, we observed less gradient as a characteristic of smoking and do not know the origin of this signature.

We note that when quantifying the test-retest reproducibility of the p_AO_2 measurements, introducing a suitable descriptive intraclass correlation coefficient (ICC) is critical (27). Due to high intersubject and intrasubject variability in p_AO_2 distributions, a random intercept mixed-effect model is utilized for defining such a coefficient (28). We constructed the model, based on the candidate fixed and random-effects that was discussed in the Methods section. After the first regression, all the insignificant covariates were removed from the model. None of the physical characteristics of the subjects (age, weight, length, BMI and TLC) show significant covariates in the linear mixed-effect model.

Short-term variation of lung function has been well documented (29–31). In this study, we extend these results to regional measurements, and also compare the variability of these regional metrics to that of standard PFT. Clinical metrics directly related to ventilation appeared to be more repeatable than the p_AO_2. This is not surprising since the p_AO_2 is a result of both variation of ventilation and perfusion. The DLCO, which might better reflect the intrasubject variation of ventilation and perfusion, had a similar repeatability to the p_AO_2. Short-term variability of the test results within a few minutes was used to compare the variability due to the testing procedure to variability due to day-to-day physiological changes within a patient. The results of these evaluations suggest that there is a sufficient test repeatability of the p_AO_2 in nonsmokers to distinguish characteristics of individual

subjects. Coupled with the differences observed between nonsmokers and healthy smokers, this suggests that imaging-based measurements of p_AO_2 could be of clinical diagnostic utility for early changes to the lung, as well as the evaluation of therapeutic response.

Acknowledgments

This work was supported by NIH grant R01-HL089064.

References

1. Benallal H, Denis C, Prieur F, Busso T. Modeling of end-tidal and arterial P_{CO2} gradient: comparison with experimental data. Med Sci Sports Exerc. 2002; 34(4):622–629. [PubMed: 11932570]
2. Jordanoglou J, Tatsis G, Danos J, Gougoulakis S, Orfanidou D, Gaga M. Alveolar Partial Pressures of Carbon-Dioxide and Oxygen Measured by a Helium Washout Technique. Thorax. 1990; 45(7): 520–524. [PubMed: 2118690]
3. Hoper J, Plasswilm L. Use of an Oxygen Multiwire Surface Electrode for the Direct Measurement of the Distribution of Alveolar pO_2 at Different Inspired Oxygen Concentrations. Clin Phys Physiol Meas. 1992; 13(3):263–271. [PubMed: 1424476]
4. Baumgardner JE, Choi IC, Vonk-Noordegraaf A, Frasch HF, Neufeld GR, Marshall BE. Sequential VA/Q distributions in the normal rabbit by micropore membrane inlet mass spectrometry. J Appl Physiol. 2000; 89(5):1699–1708. [PubMed: 11053316]
5. Eberle B, Weiler N, Markstaller K, Kauczor HU, Deninger A, Ebert M, Grossmann T, Heil W, Lauer LO, Roberts TPL, Schreiber WG, Surkau R, Dick WF, Otten EW, Thelen M. Analysis of intrapulmonary O-2 concentration by NIR imaging of inhaled hyperpolarized helium-3. J Appl Physiol. 1999; 87(6):2043–2052. [PubMed: 10601148]
6. Deninger AJ, Eberle B, Ebert M, Grossmann T, Heil W, Kauczor HU, Lauer L, Markstaller K, Otten E, Schmiedeskamp J, Schreiber W, Surkau R, Thelen M, Weiler N. Quantification of regional intrapulmonary oxygen partial pressure evolution during apnea by He-3 MRI. J Magn Reson. 1999; 141(2):207–216. [PubMed: 10579944]
7. Deninger AJ, Eberle B, Ebert M, Grossmann T, Hanisch G, Heil W, Kauczor HU, Markstaller K, Otten E, Schreiber W, Surkau R, Weiler N. ^{3}He-MRI-based measurements of intrapulmonary p_{O2} and its time course during apnea in healthy volunteers: first results, reproducibility, and technical limitations. NMR Biomed. 2000; 13(4):194–201. [PubMed: 10867696]
8. Deninger AJ, Eberle B, Bermuth J, Escat B, Markstaller K, Schmiedeskamp J, Schreiber WG, Surkau R, Otten E, Kauczor HU. Assessment of a single-acquisition imaging sequence for oxygen-sensitive ^{3}He MRI. Magn Reson Med. 2002; 47(1):105–114. [PubMed: 11754449]
9. Fischer MC, Spector ZZ, Ishii M, Yu J, Emami K, Itkin M, Rizi R. Single-acquisition sequence for the measurement of oxygen partial pressure by hyperpolarized gas MRI. Magn Reson Med. 2004; 52(4):766–773. [PubMed: 15389934]
10. Fischer MC, Kadlecek S, Yu JS, Ishii M, Emami K, Vahdat V, Lipson DA, Rizi RR. Measurements of regional alveolar oxygen pressure using hyperpolarized ^{3}He MRI. Acad Radiol. 2005; 12(11):1430–1439. [PubMed: 16253855]
11. Ishii M, Fischer MC, Emami K, Alavi A, Spector ZZ, Yu JS, Baumgartner JE, Itkin M, Kadlecek SJ, Zhu JL, Bono M, Gefter WB, Lipson DA, Shrager JB, Rizi RR. Hyperpolarized helium-3 MR imaging of pulmonary function. Radiol Clin N Am. 2005; 43(1):235–246. [PubMed: 15693659]
12. Wild JM, Fichele S, Woodhouse N, Paley MNJ, Kasuboski L, van Beek JR. 3D volume-localized pO_2 measurement in the human lung with ^{3}He MRI. Magn Reson Med. 2005; 53(5):1055–1064. [PubMed: 15844148]
13. Yu J, Rajaei S, Ishii M, Law M, Emami K, Woodburn JM, Kadlecek S, Vahdat V, Rizi RR. Measurement of pulmonary partial pressure of oxygen and oxygen depletion rate with hyperpolarized helium-3 MRI: A preliminary reproducibility study on pig model. Acad Radiol. 2008; 15(6):702–712. [PubMed: 18486007]

14. Cieslar K, Stupar V, Canet-Soulas E, Gaillard S, Cremillieux Y. Alveolar oxygen partial pressure and oxygen depletion rate mapping in rats using He-3 ventilation imaging. Magn Reson Med. 2007; 57(2):423–430. [PubMed: 17152086]
15. Kraayvanger RJ, Bidinosti CP, Dominguez-Viqueira W, Parra-Robles J, Fox M, Lam WW, Santyr GE. Measurement of Alveolar Oxygen Partial Pressure in the Rat Lung Using Carr-Purcell-Meiboom-Gill Spin-Spin Relaxation Times of Hyperpolarized ^{3}He and ^{129}Xe at 74 mT. Magn Reson Med. 2010; 64(5):1484–1490. [PubMed: 20593371]
16. Ouriadov AV, Lam WW, Santyr GE. Rapid 3-D mapping of hyperpolarized He-3 spin-lattice relaxation times using variable flip angle gradient echo imaging with application to alveolar oxygen partial pressure measurement in rat lungs. Magn Reson Mat Phys Biol Med. 2009; 22(5): 309–318.
17. Miller GW, Mugler JP III, Altes TA, Cai J, Mata JF, de lange EE, Tobias WA, Cates GD, Brookeman JR. A Short-Breath-Hold Technique for Lung pO_2 Mapping With ^{3}He MRI. Magn Reson Med. 2010; 63(1):127–136. [PubMed: 19918891]
18. Hamedani H, Kadlecek SJ, Emami K, Kuzma NN, Xu Y, Xin Y, Mongkolwisetwara P, Rajaei J, Barulic A, Wilson Miller G, Rossman M, Ishii M, Rizi RR. A multislice single breath-hold scheme for imaging alveolar oxygen tension in humans. Magnetic Resonance in Medicicne. 2011
19. Klaus Kurt Gast AB, Annette Herweling, Wolfgang Günter Schreiber, Jörg Schmiedeskamp, Eckhard Mayer, Claus Peter Heussel, Klaus Markstaller, Hans-Ulrich Kauczor, Balthasar Eberle. Oxygen-sensitive 3He-MRI in bronchiolitis obliterans after lung transplantation. Eur Radiol. 2008 Mar; 18(3):530–537. [PubMed: 17926041]
20. Jalali A, Ishii M, Edvinsson JM, Guan L, Itkin M, Lipson DA, Baumgardner JE, Rizi RR. Detection of simulated pulmonary embolism in a porcine model using hyperpolarized ^{3}He MRI. Magn Reson Med. 2004; 51(2):291–298. [PubMed: 14755654]
21. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J. Standardisation of spirometry. Eur Resp J. 2005; 26(2):319–338.
22. Saam B, Happer W, Middleton H. Nuclear Relaxation of ^{3}He in the Presence of O_2. Phys Rev A. 1995; 52(1):862–865. [PubMed: 9912313]
23. Z. Y, Z. H. Lung Image Volume Registration Using the Affine Transformation and Demons Algorithm. Bioinformatics and Biomedical Engineering (iCBBE). 2010 Jun.
24. Akaike, H. Likelihood and the Bayes procedure (with discussion). In: Bernardo, JM.; De Groot, MH.; Lindley, DV.; Smith, AFM., editors. Bayesian Statistics. Valencia, Spain: University Press; 1980. p. 1143-203.
25. A. HD. Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. Journal of the American Statistical Association. 1977 Jun; 72(358):320–338.
26. Rosner, B. 6th ed. Belmont, CA: Duxbury Press; 2006. Fundamentals of Biostatistics.
27. Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. Statistics in Medicine. 1997; 16(7):821–822. [PubMed: 9131768]
28. Donnar A, Koval J. The Estimation of Intraclass Correlation in the Analysis of Family Data. Biometrics. 1980 Mar; 36(1):19–25. [PubMed: 7370372]
29. Borsboom GJJM, van Pelt W, van Houwelingen HC, van Vianen BG, Schouten JP, Quanjer PH. Diurnal variation in lung function in subgroups from two Dutch populations - Consequences for longitudinal analysis. Am J Respir Crit Care Med. 1999; 159(4):1163–1171. [PubMed: 10194161]
30. Tweeddale PM, Alexander F, McHardy GJR. Short-Term Variability In Fev1 And Bronchodilator Responsiveness In Patients With Obstructive Ventilatory Defects. Thorax. 1987; 42(7):487–490. [PubMed: 3438892]
31. Rozas CJ, Goldman AL. Daily Spirometric Variability - Normal Subjects And Subjects With Chronic-Bronchitis With And Without Air-Flow Obstruction. Arch Intern Med. 1982; 142(7): 1287–1291. [PubMed: 7092446]

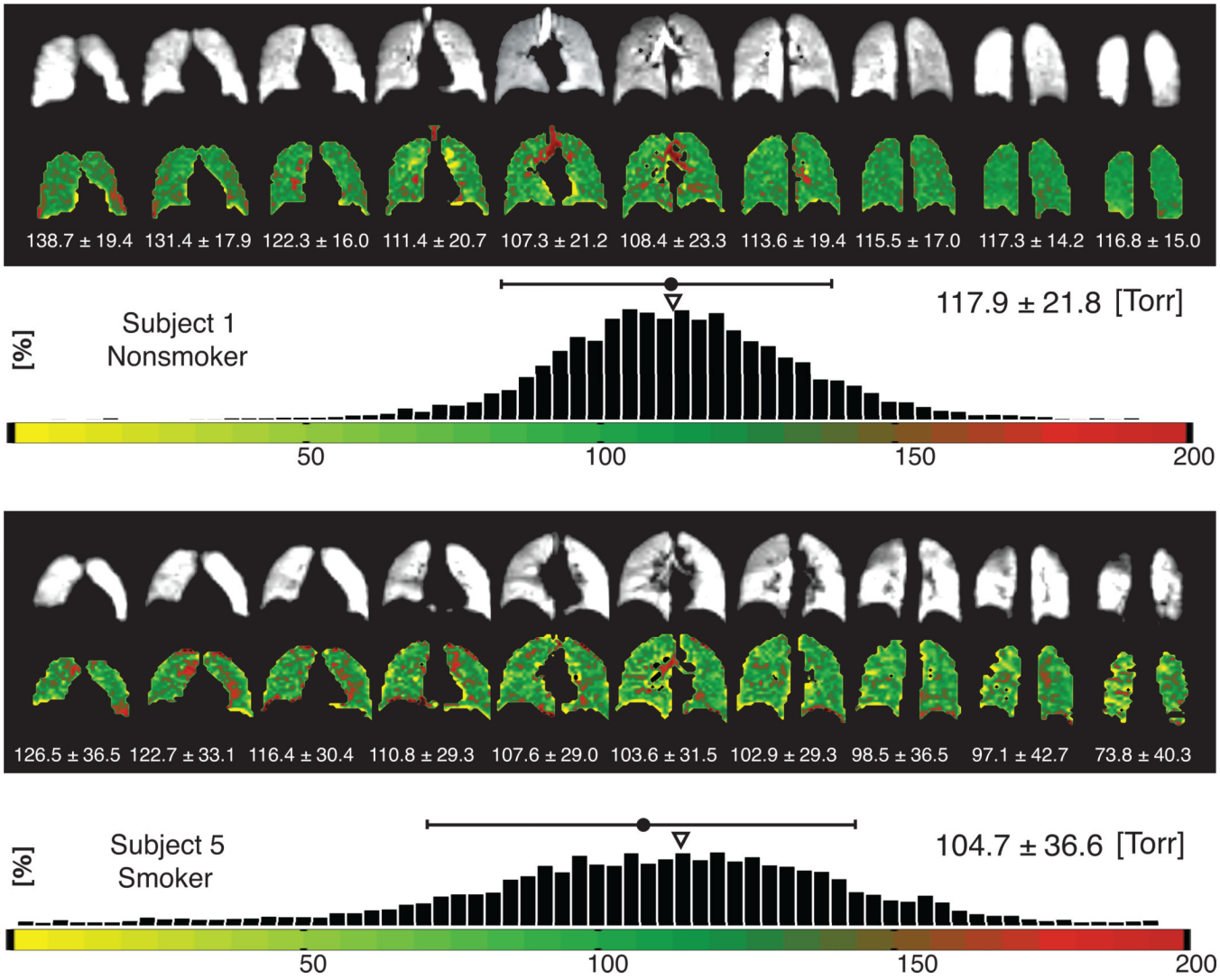

Figure 1.

Whole-lung signal intensity (B/W) and p_AO_2 maps for representative nonsmoker (top) and smoker (bottom) subjects and the associated $p_AO_2$histograms for all the voxels. The mean and standard deviation is listed above the histograms. The small triangle over the histograms shows the median of the distribution.

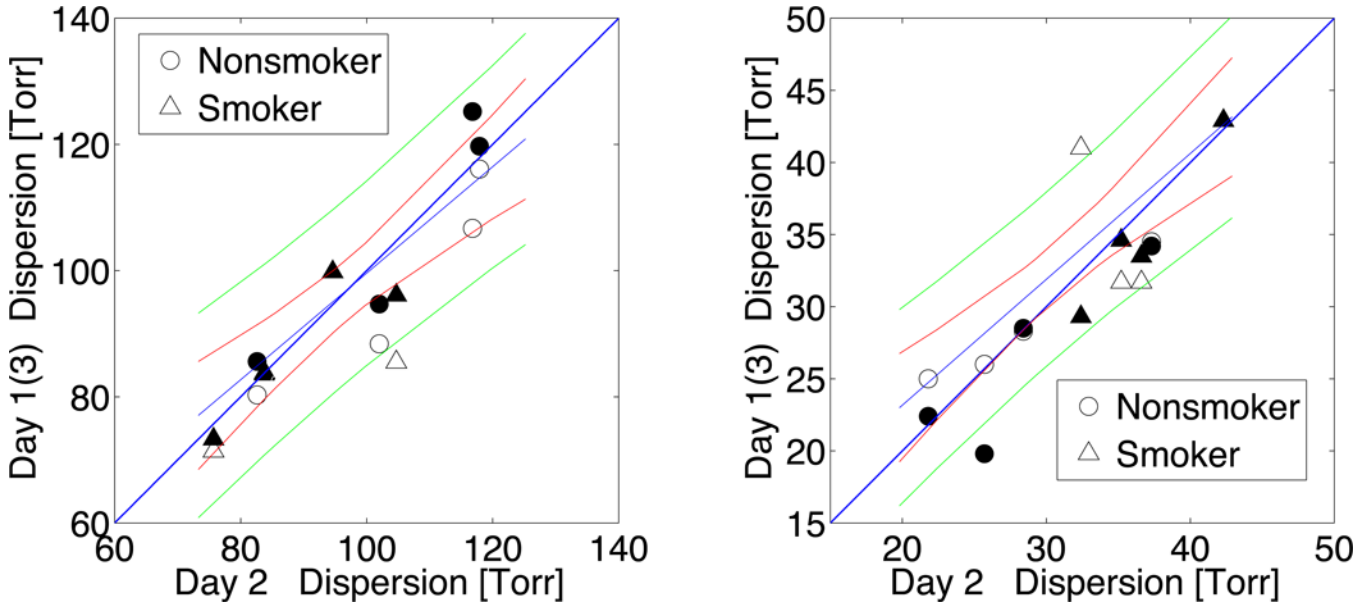

Figure 2.

Day-to-day correlation between whole-lung p_AO_2 means and standard deviations (dispersion) from the same subject. All the day's measurements is shown in one figure; hollow-markers indicate day1 vs. day2 and the solid-markers indicate day2 vs. day3 measurements.

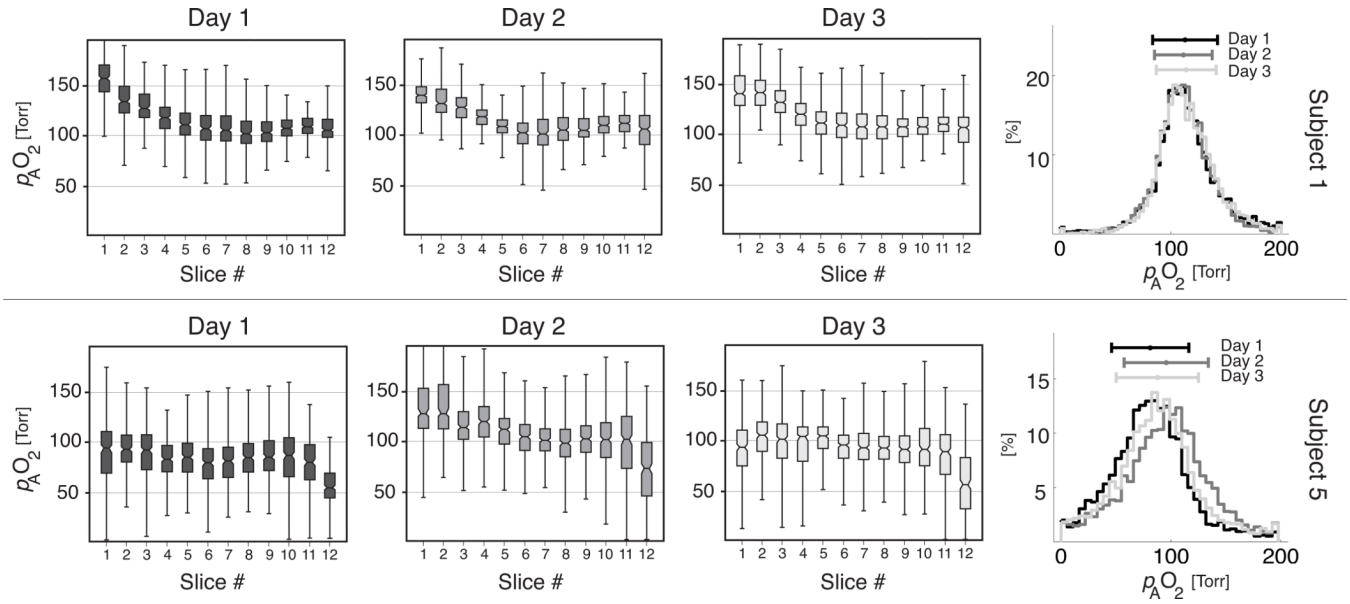

Figure 3.

Derived p_AO_2 measurements by slice for representative nonsmoker (top) and smoker (bottom) subjects by slice on each study day (left panel) and same-subjects whole-lung histograms from each study day (right panel).

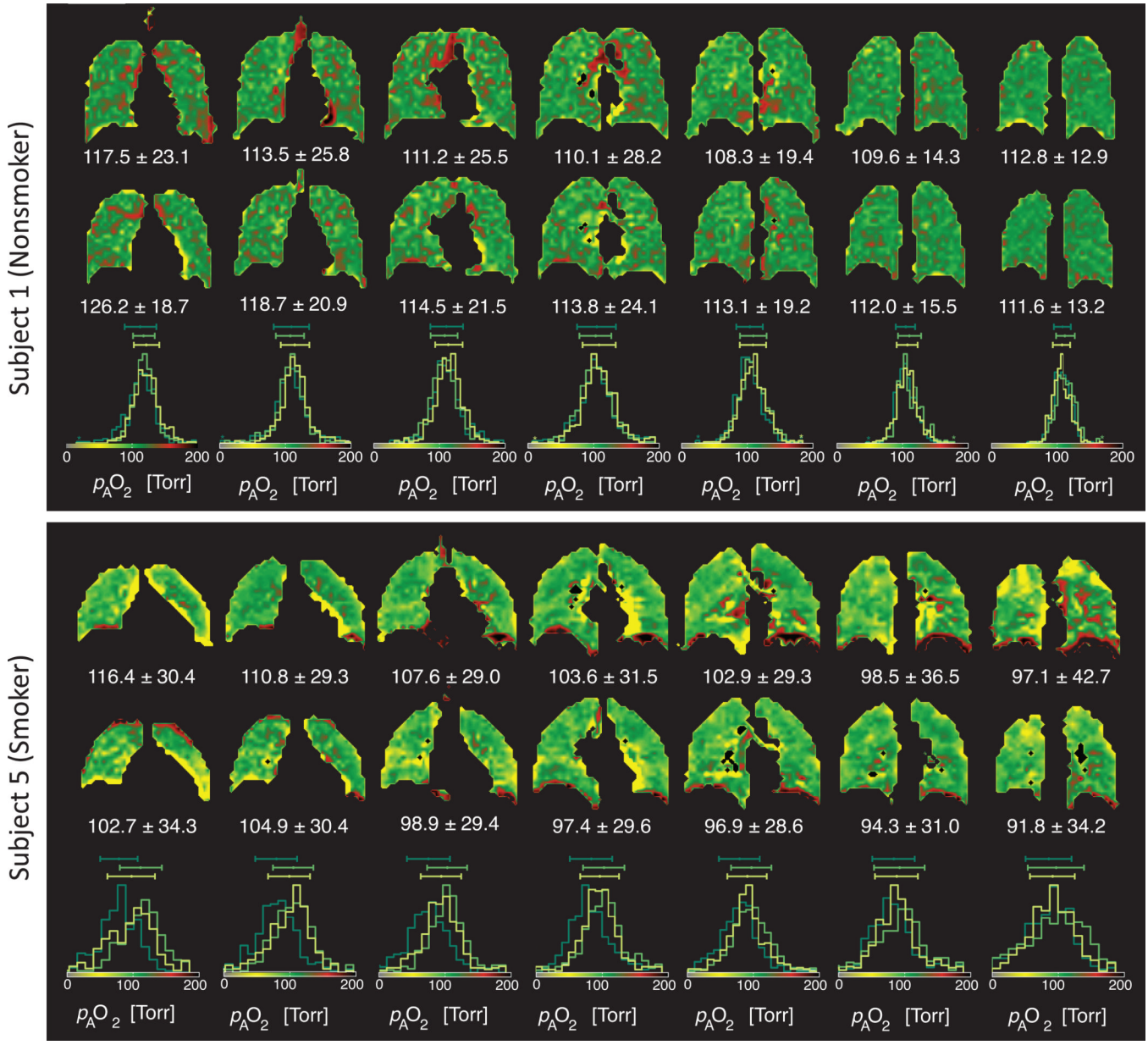

Figure 4.

Day-to-day comparison of same-subject p_AO_2 maps (Torr) in the representative nonsmoker (top) and smoker subject (bottom). Slices are shown in the anterior to posterior direction. Seven middle slices are shown (##4–10) from days 2 and 3. along with the corresponding overlaid histograms from all three days.

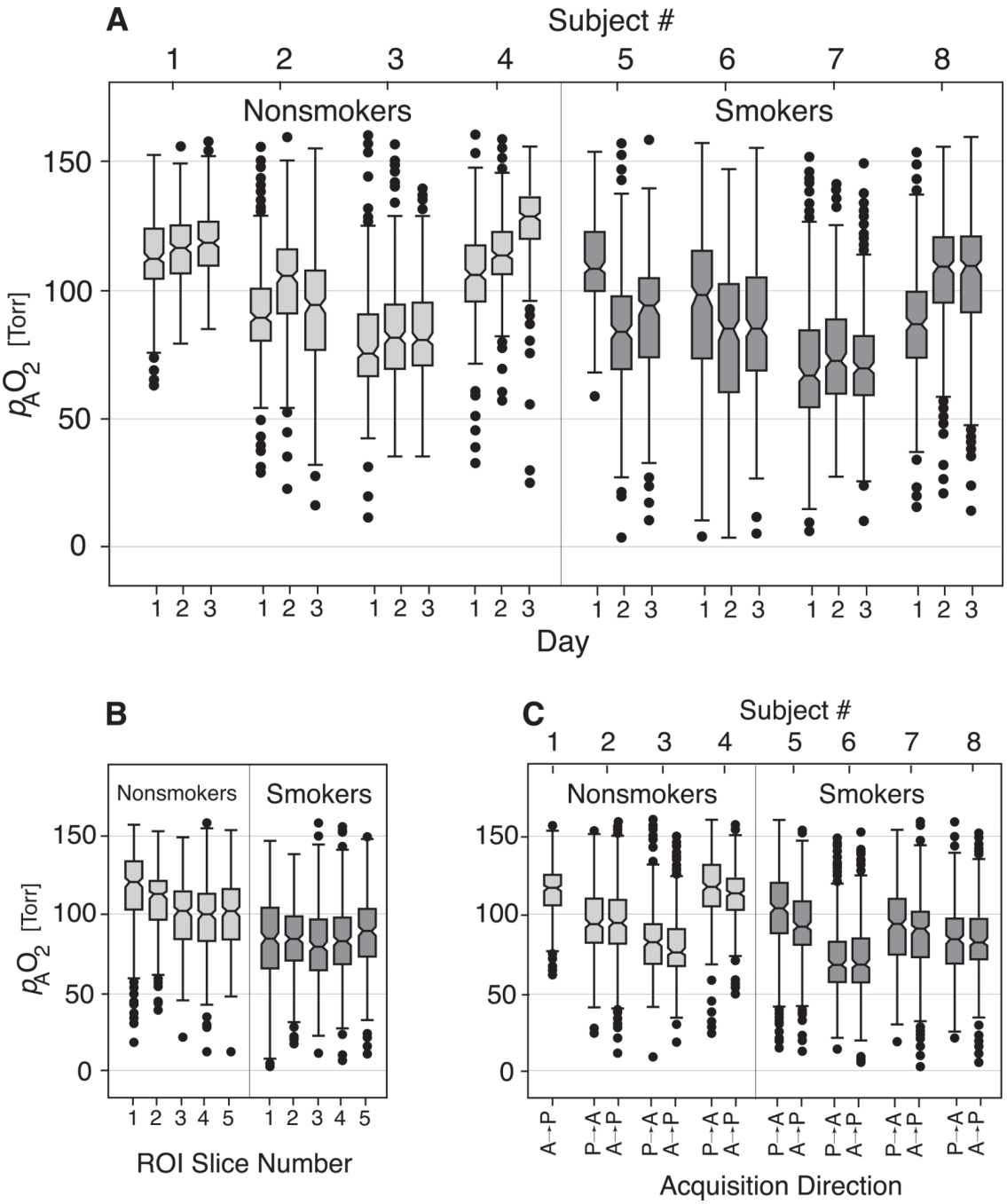

Figure 5.
Summary box plots depicting p_AO_2 variation with day (A), slice (B), and acquisition direction (C).

Table 1

Summary of subject demographic characteristics

Subject		Age [Yrs]	Sex	Weight [Lb]	BMI† [kg/m^2]	Pack Years	TLC†† [L]	P_AO_2 (mean ± standard deviation) [Torr]		
								Day 1	Day 2	Day 3
Nonsmokers	1	56	F	150	23.4	-	5.9	116.1 ± 25.0	117.9 ± 21.8	119.7 ± 22.4
	2	56	M	185	28.5	-	7.4	88.4 ± 28.3	102.0 ± 28.4	94.7 ± 28.5
	3	57	F	180	29.4	-	4.6	80.3 ± 34.5	82.6 ±37.3	85.6 ± 34.2
	4	53	F	170	23.7	-	7.1	106.7 ± 26.0	116.8 ± 25.7	125.2 ± 19.8
	ave ± std	56 ± 1.7	-	171 ± 15	26.3 ± 3.1	-	6.3 ± 1.3			
smokers	5	47	M	190	26.2	30	8.5	85.5 ± 31.7	104.7 ± 36.6	96.1 ±33.5
	6	52	M	230	32.2	35	5.5	71.4 ± 41.0	75.7 ± 32.4	73.3 ± 29.3
	7*	63	F	115	21.0	45	4.4	-	94.6 ±42.3	99.8 ± 42.9
	8	47	M	180	26.5	25	6.4	83.8 ± 31.7	83.8 ±35.2	83.6 ± 34.6
	ave ± std	52 ± 7.5	-	179 ± 48	26.5 ± 4.6	34 ± 9	6.2 ± 1.7			
	ave ± std	**54 ± 5.4**		**175 ± 33**	**26.4 ± 3.6**	**34 ± 9**	**6.2 ± 1.4**			

† Body Mass Index: weight [kg] / length^2 [m^2]

†† Average Total Lung Capacity

* Subject 7 was not able to inhale the gas contecnt correctly and failed the breath-hold in Day 1.

Table 2

Summarized pulmonary function test results from all study subjects and comparison with p_AO_2 measurements.

		PFT Results in 3 Days															*	p_AO_2 in 3 Days	
		FVC†			FEV_1†			$FEF_{25-75\%}$†			TLC†			DLCO†				μp_AO_2	σp_AO_2
		Mean	SD	CV††	Mean	SD	CV††	Mean	SD	CV††	Mean	SD	CV††	Mean	SD	CV††		CV††	CV††
Subject		[L]		[%]	[L]		[%]	[L/s]		[%]	[L]		[%]	[ml/min/mmHg/L]		[%]		[%]	[%]
Nonsmokers	1	3.88	0.09	2.23	3.29	0.06	1.85	3.90	0.03	0.64	5.78	0.11	1.85	26.16	1.46	5.57		1.53	7.37
	2	5.03	0.03	0.64	4.03	0.04	0.87	3.75	0.32	8.58	7.26	0.11	1.53	31.74	0.74	2.34		7.16	0.35
	3	2.68	0.05	1.97	2.15	0.05	2.14	2.09	0.21	10.18	4.45	0.19	4.27	22.51	1.42	6.30		3.21	4.84
	4	4.90	0.08	1.66	4.03	0.07	1.65	4.33	0.17	3.95	6.89	0.40	5.87	25.71	1.78	6.91		7.97	14.67
CVave±std		**1.63 ± 0.7**			**1.63 ± 0.5**			**5.84 ± 4.4**			**3.38 ± 2.1**			**5.28 ± 2.0**			μ	**4.97 ± 3.1**	
																	σ	**6.81 ± 6.0**	
Smokers	5	6.42	0.16	2.54	4.44	0.08	1.76	2.88	0.12	4.09	8.38	0.14	1.69	23.81	0.87	3.65		10.08	7.30
	6	3.86	0.10	2.70	3.08	0.09	3.02	2.99	0.25	8.21	5.88	0.41	6.91	29.81	2.54	8.54		2.93	17.71
	7	2.50	0.02	0.61	2.05	0.02	1.01	2.17	0.04	2.01	4.40	0.04	0.82	10.93	1.18	10.8		3.78	1.00
	8	4.55	0.02	0.51	3.60	0.09	2.41	3.32	0.28	8.37	6.23	0.22	3.58	27.56	1.50	5.44		0.14	5.53
CVave±std		**1.59 ± 1.2**			**2.05 ± 0.9**			**5.67 ± 3.1**			**3.25 ± 2.1**			**5.84 ± 2.5**			μ	**4.23 ± 4.2**	
																	σ	**7.88 ± 7.1**	

**		FVC†	FEV_1†	$FEF_{25-75\%}$†	TLC†	DLCO†
t-test	μp_AO_2	0.02	0.04	0.74	0.35	0.44
	σp_AO_2	0.02	0.02	0.55	0.07	0.74

† FVC: Forced Vital Capacity; FEV_1: Forced Expiratory Flow in 1 Second; FEF: Forced Expiratory Flow; TLC: Total Lung Capacity; DLCO: Diffusing Capacity

†† Coefficient of Variation

* The means and standard deviations are reported in Table 1

** The results of paired t-test between coefficient of variations of imaging results and PFT properties

Table 3

Mixed-effect model coefficients, confidence intervals and significance for multiple measurements across a two-week period.

		Coefficient [Torr]	Standard Error [Torr]	*P*-value	CI† [Torr]
Fixed Effects	**Constant**	121.556	7.543	<0.001	106.773,136.340)
	Smoker	−33.574	9.862	0.001	−(52.903,14.243)
	Binned Slice	−10.562	2.58	<0.001	−(15.617,5.506)
	Binned Slice^2	1.139	0.4138	0.006	(0.328,1.950)
	B. Slice x Smoker	4.992	0.992	<0.001	(3.047,6.936)
		Estimate SD [Torr]			
Rand. Effects	**Subject**	13.127	3.874		(7.361,23.408)
	Slice	3.565	0.755		(2.353,5.401)
	Voxel	11.119	0.451		(10.269,12.039)
Residuals	**Normal**	14.598	0.269		(14.079,15.136)
	Smoker	18.839	0.408		(18.056,19.656)

† The numbers in parantheses are 95% Confidence Intervals.

Table 4

Mixed-effect model coefficients, confidence intervals and significance for multiple measurements for short-term test-retest measurements.

		Coefficient [Torr]	Standard Error [Torr]	*P*-value	CI† [Torr]
Fixed Effects	**Constant**	142.537	17.701	<0.001	107.842,177.232)
	Smoker	−23.375	10.974	0.033	−(44.884,1.866)
	Binned Slice	−12.828	3.378	0.004	−(19.449,-6.207)
	Binned Slice^2	1.583	0.13	0.004	(0.520,2.645)
	B. Slice x Smoker	4.302	1.299	0.0011	(1.756,6.848)
		Estimate SD [Torr]			
Rand. Effects	**Subject**	14.24	3.686		(8.574,23.651)
	Slice	4.833	0.916		(3.333,7.008)
	Voxel	13.65	0.527		(12.654,14.723)
Residuals	**Normal**	12.945	0.381		(12.218,13.714)
	Smoker	15.682	0.538		(14.661,16.775)

†The numbers in parantheses are 95% Confidence Intervals.

Table 5

The difference between the short- and long-term variability.

Difference (Long-term - Short-term)	Estimate	Bias	*Bootstarp* Standard Error	CI†
Nonsmoker	−0.151	0.021	0.021	−(0.193 , 0.110)
Smoker	−0.198	0.02	0.253	−(0.248 , 0.148)