ARTICLE IN PRESS

Original Investigation

Influence of Inspiratory/Expiratory CT Registration on Quantitative Air Trapping

Oliver Weinheimer, Benjamin A. Hoff, Aleksa B. Fortuna, Antonio Fernández-Baldera, Philip Konietzke, Mark O. Wielpütz, Terry E. Robinson, Craig J. Galbán, PhD

Abbreviations

CT Computed Tomography
QAT quantitative air trapping
mHU mean Hounsfield Units
DC Dice Coefficient
CF Cystic Fibrosis
HU Hounsfield Units
3D 3-dimensional
QCT quantitative CT
AT air trapping
SAD small airways disease
V_I total lung volume at inspiration
V_E total lung volume at expiration
V_{HI} volume of healthy parenchyma at inspiration
V_{HE} volume of healthy parenchyma at expiration

Rationale and Objectives: The aim of this study was to assess variability in quantitative air trapping (QAT) measurements derived from spatially aligned expiration CT scans.

Materials and Methods: Sixty-four paired CT examinations, from 16 school-age cystic fibrosis subjects examined at four separate time intervals, were used in this study. For each pair, visually inspected lobe segmentation maps were generated and expiration CT data were registered to the inspiration CT frame. Measurements of QAT, the percentage of voxels on the expiration CT scan below a set threshold were calculated for each lobe and whole-lung from the registered expiration CT and compared to the true values from the unregistered data.

Results: A mathematical model, which simulates the effect of variable regions of lung deformation on QAT values calculated from aligned to those from unaligned data, showed the potential for large bias. Assessment of experimental QAT measurements using Bland-Altman plots corroborated the model simulations, demonstrating biases greater than 5% when QAT was approximately 40% of lung volume. These biases were removed when calculating QAT from aligned expiration CT data using the determinant of the Jacobian matrix. We found, by Dice coefficient analysis, good agreement between aligned expiration and inspiration segmentation maps for the whole-lung and all but one lobe (Dice coefficient > 0.9), with only the lingula generating a value below 0.9 (mean and standard deviation of 0.85 ± 0.06).

Conclusion: The subtle and predictable variability in corrected QAT observed in this study suggests that image registration is reliable in preserving the accuracy of the quantitative metrics.

Key Words: Cystic Fibrosis; Computed Tomography; Image registration; Multiparametric analysis.

Acad Radiol 2018; ■:1–13

From the Department of Diagnostic and Interventional Radiology, University Hospital of Heidelberg, 69120 Heidelberg, Germany (O.W., P.K., M.O.W.); Translational Lung Research Center, Heidelberg (TLRC), German Lung Research Center (DZL), 69120 Heidelberg, Germany (O.W., P.K., M.O.W.); Department of Radiology, University of Michigan, Ann Arbor, MI 48109 (B.A.H., A.B.F., A.F.-B., C.J.G.); Center of Excellence in Pulmonary Biology, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94304 (T.E.R.). Received September 17, 2018; revised October 25, 2018; accepted November 3, 2018. Conflict of Interest Disclosure: BAH, ABF, AFB, PK, TER, and CJG have no relevant conflicts of interest to disclose. MOW acts as a consultant for Boehringer Ingelheim and Vertex, and has received funding from Boehringer Ingelheim and from the German Cystic Fibrosis Foundation (Mukoviszidose e.V.). OW has a financial interest in the segmentation technology licensed to Imbio, L.L.C. Funding: This work was supported by the US National Institutes of Health research grants R44HL118837 and R44HL140894. **Address correspondence to:** C.J.B. e-mail: cgalban@med.umich.edu

https://doi.org/10.1016/j.acra.2018.11.001

V_{AT}
volume of air trapping

QAT_E
QAT is defined from the expiration CT

QAT_{E2I}
QAT is defined from the expiration CT scan spatially aligned to the inspiration geometric frame where E2I means expiration lungs-to-inspiration frame

V_{E2I}
expiration lung volume mapped to the inspiration space

FEV1
forced expiratory volume at one second

TLC
total lung capacity

RV
residual volume

InsSeg
lobe segmentation map for inspiration CT scan

ExpSeg
lobe segmentation map for expiration CT scan

Exp
expiration CT scan

Ins
inspiration CT scan

ExpR
expiration CT scan aligned to the inspiration geometric frame

SegExpR
expiration lobe segmentation map aligned to the inspiration geometric frame

(T)
threshold

(J)
Jacobian determinant map

BA plots
Bland-Altman

RUL
right upper lobe

RML
right middle lobe

RLL
right lower lobe

LUL
left upper lobe

LLi
lingula

LLL
left lower lobe

ARTICLE IN PRESS

Academic Radiology, Vol ■, No ■■, ■■ 2018 INFLUENCE OF INSPIRATORY/EXPIRATORY CT REGISTRATION

(QAT_1) QAT values calculated using Method 1

(QAT_2) QAT values calculated using Method 2

(QAT_3) QAT values calculated using Method 3

(QAT_4) QAT values calculated using Method 4

INTRODUCTION

A variety of 3-dimensional (3D) techniques that demonstrate prognostication in lung diseases have been developed for quantitative CT (QCT) (1,2). The most extensively used method is the quantification of low attenuation areas using a threshold-based approach. When applying a threshold of −950 Hounsfield Units (HU) to CT scans acquired at inspiration (i.e. full-inflation), this quantitative index, presented as the relative volume of lung parenchyma, has been pathologically validated as a measure of emphysema (3–6). In a similar fashion, threshold-based techniques applied to expiratory CT scans have been used to assess the extent of air trapping (AT), a hallmark of small airways disease (7–9). The high air-tissue contrast on the inspiration CT scan has also been exploited to develop methods for airway and vessel measurements, as well as lobe segmentation (10–14). Combined, all these approaches provide detailed quantitative information on airway and vessel remodeling and alterations in local parenchyma but are relegated to the scan acquisition volume (i.e. geometric space) from which the CT scan was acquired.

The wide-spread availability of image registration algorithms has allowed for the rapid advancement of new QCT analytical techniques (15, 16). These registration-based methods have been shown to generate ventilation maps (17, 18), improve air trapping measurements (19, 20), correct differences in inflation levels for serial quantitative emphysema measurements (21, 22), and provide additional phenotyping of obstructive diseases as demonstrated in COPD and bronchiolitis obliterans syndrome (23–26). Due to wide acceptance of these techniques in pulmonary medicine, registration-based algorithms are commercially available in software packages for quantitative analysis of lung CT scans. This has further broadened the availability of these techniques to the clinical community as well as their use in analyzing CT data from large clinical trials (e.g. COPDGene (27) and SPIROMICS (28)).

Registration-based techniques applied to serial thoracic CT examinations are advancing the clinical care of patients suffering from lung disease. Although extensive research has been devoted to improving the accuracy in these algorithms (29), little has been done to evaluate the effect of image registration on QCT metrics. In this study, we demonstrated sources of variability in QCT metrics as a consequence of image registration and propose a simple strategy to help correct for the deviation observed in QCT values.

MATERIALS AND METHODS

Deviation of Aligned QAT

Here we derived a mathematical expression that simulates the variability of quantitative air trapping (QAT) measurements, defined below, determined from spatially aligned and nonaligned expiration CT scans. As illustrated in Figure 1, the lungs are assumed to consist of a two phase system: healthy and air-trapped parenchyma. The healthy parenchyma is assumed to undergo deformation from expiration to inspiration. In contrast, air-trapped parenchyma is assumed to have a constant volume at both lung phases, i.e. incompressible and cannot be ventilated beyond the inflation levels. The volumes of the two phases at each inflation level are represented as:

$$V_I = V_{HI} + V_{AT} \text{ and } V_E = V_{HE} + V_{AT}, \quad (1)$$

where V_I and V_E define the total lung volume at inspiration and expiration, respectively, and V_{HI}, V_{HE}, and V_{AT} indicate the volume of healthy parenchyma at inspiration and expiration, respectively, and the volume of air trapping. The accepted convention for QAT is to normalize the volume of air trapping to the total lung volume on the expiration CT scan (27). Using this approach for calculating QAT, we define QAT using the nonaligned and aligned expiration CT scans as:

$$QAT_E = V_{AT}/V_E \text{ and } QAT_{E2I} = V_{AT}/V_{E2I}, \quad (2)$$

where QAT_{E2I} indicates that QAT is determined from the expiration CT scan spatially aligned to the inspiration geometric frame (E2I means expiration lungs-to-inspiration frame) and V_{E2I} is the expiration lung volume mapped to the inspiration space ($V_{E2I} = V_I$). Using the equations presented in (1) and (2), QAT_{E2I} can be expressed as a function of QAT_E and the volume of healthy parenchyma at inspiration and expiration:

$$QAT_{E2I} = \frac{\alpha\, QAT_E}{1 - QAT_E(1-\alpha)}, \quad (3)$$

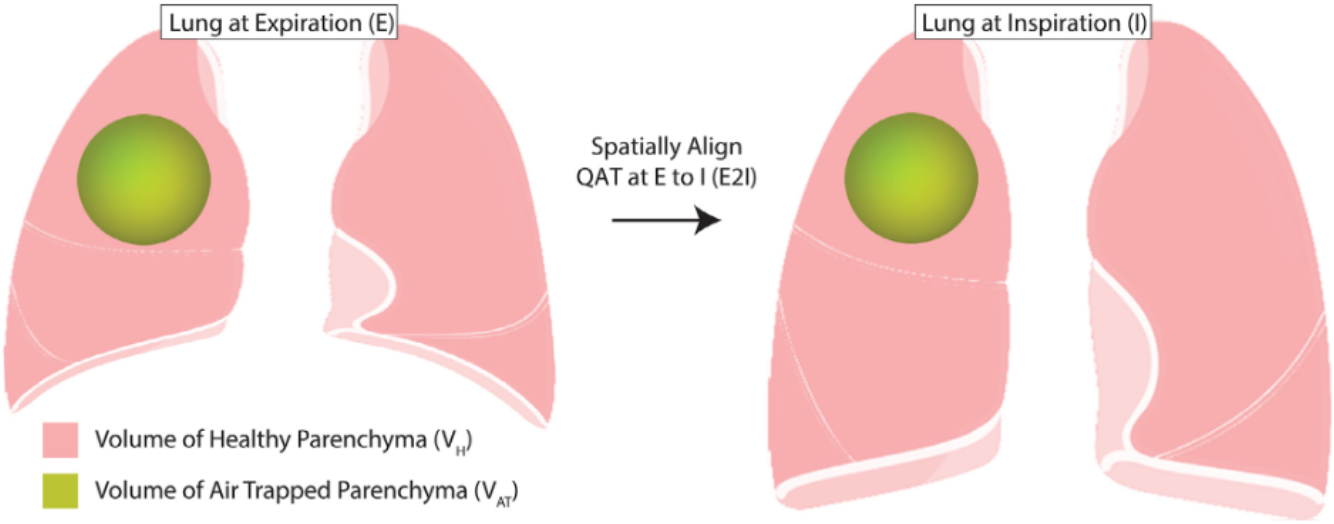

Figure 1. Illustration of the effect of lung inflation level on volume changes in healthy and air-trapped parenchyma. For our mathematical derivation we assumed that the lung consists of two-phases: healthy (H) and air-trapped (AT) parenchyma, depicted as sphere in right upper lobe of lungs. The direction of registration is the expiration (E) CT scan mapped to the inspiration (I) CT frame (E2I). The spatial deformation from expiration to inspiration is assumed to only occur within the healthy parenchyma, with the air-trapped volume assumed constant.

where α is the ratio V_{HE}/V_{HI}. The full derivation is provided in Supplemental section.

Subjects

CF subjects were school-age children (mean ± standard deviation of 11.1 ± 2.4 years) accrued as part of the Novartis/CF Foundation Therapeutics prospective 2-year natural history study (30). From this cohort, we analyzed CT data acquired from 16 Stanford CF subjects (Table 1) with baseline and follow-up examinations at 3, 12, and 24 months using the identical scanner and protocol. CF subjects were extensively characterized at baseline based on age, gender, height, weight, body mass index, forced expiratory volume at 1 second (FEV1 and FEV1 percent predicted), total lung capacity (TLC and TLC percent predicted) and residual volume (RV and RV percent predicted).

TABLE 1. Baseline Characteristics

Parameter	
Number	16
Age (years)	11.1±2.4
Gender (male/female)	9/7
Weight (kg)	39.2±10.3
Height (cm)	144±14.1
BMI (kg/m^2)	18.5±2.2
FEV1 (L/s)	2.3±0.6
FEV1 (% predicted)	106±11.9
TLC (L)	3.7±0.9
TLC (% predicted)	108±11.2
RV (L)	0.9±0.3
RV (% predicted)	131±38.4

Note: All data is presented as the mean ± standard deviation. BMI is the body mass index. FEV1 is the forced expiratory volume at 1 sec. TLC is the total lung capacity. RV is the residual volume.

CT Acquisition

Paired inspiratory and expiratory multidetector CT scans (**Ins** and **Exp**, respectively) were acquired on a Siemens Sensation 64 scanner (Siemens Healthcare AG, Forchheim, Germany). All examinations were performed in the supine position utilizing a spirometer-controlled imaging technique as previously described (8). For all subjects, an identical low dose spiral CT scanning protocol was utilized with 100 kVp, 30−50 mAs, pitch of 1.0 (inspiratory) and 1.2 (expiratory), collimation was 0.6 mm and reconstructed slice thickness was 0.6 mm with 50% overlap in lung and soft-tissue kernels (b60f and b30f, respectively) (9, 31, 32). The calculated total combined effective dose for the inspiratory and expiratory chest CT scans from Packard Children's Hospital at Stanford was 1.35 mSv. This corresponded to an estimated risk of developing cancer of ~ 0.0135% (33, 34), meaning that for every 1000 people exposed to comparable radiation exposures, 999.87 people on average would not develop cancer due to this exposure. Quality assurances were determined visually by a single reader with greater than 25 years of experience in pediatric chest radiology for adequate inspiration, absence of significant motion artifacts, and inclusion of all parts of the chest. Variables for 3D data have been bolded (e.g. **Exp** and **Ins**) so that they may be differentiated from whole-volume scalar quantities (e.g. QAT).

Image Segmentation and Registration

The lobes and lungs from both CT scans were segmented from the surrounding anatomy (i.e., bronchus, heart, and

chest wall) using a segmentation algorithm developed by investigators MOW and OW at University Hospital of Heidelberg (YACTA, version 2.7) as previously described (3, 11–13, 31). All lobe segmentation maps for all time-points and for both inspiratory and expiratory CT (**InsSeg** and **ExpSeg**, respectively) were visually verified and manually corrected by a single reader ($16 \times 4 \times 2 = 128$ segmentation maps). Following generation of the segmentation maps, the expiration CT scan (**Exp**) was registered to the inspiration CT scan (**Ins**) for each patient and time point ($16 \times 4 = 64$ registrations). All registrations were performed using Elastix (version 4.8), an open-source deformable image registration library (16, 35). This algorithm iteratively optimizes the solution using mutual information with a bending energy penalty as the objective function. The optimized transformation matrix was used to align the original expiratory CT scan and corresponding segmentation maps to the inspiratory geometric frame using nearest-neighbor interpolation (**ExpR** and **SegExpR**, representative).

Quantitative Air Trapping and Local Deformation Correction

QAT was determined for both **Exp** and **ExpR** using a threshold-based approach. A slight variation on a previously reported algorithm was used to generate subject-specific thresholds for detecting regions of parenchyma with mild-to-severe air trapping (originally defined QAT_{A1}) (8). In brief, a 3^3, i.e. subvolume of 27 voxels, median filter was applied to **Ins** and **Exp** immediately prior to AT classification (36). The whole lungs were then segmented from the individual filtered CT data and voxels with HU values > 0 were excluded. The 50th percentile for **Ins** (Y), the 90th percentile for **Ins** (X) and difference in the 90th percentile values in **Ins** and **Exp** (D) were determined. These values in HU were used in the following expression to determine a subject-specific threshold (T) for air trapping, $T = X - (1-D/343)^{\star}(X-Y)/3$ (8). An air trapping map (**AT**) was generated by classifying all **Exp** voxels with HU values <T as 1 and the remaining voxels as 0. QAT was calculated from **AT** by summing the binary value of all voxels and normalized to the total number of voxels within the segmentation map (i.e. whole-lung and individual lobes). Using the same threshold (T), the process was repeated for **ExpR** to calculate the respective QAT. In addition to QAT, the mean HU was determined for the whole-lung and individual lobes. For reference, air and water attenuation values are −1000 and 0 HU, respectively.

To correct for differences in lung inflation levels between **ExpR**, which has the same volume as **Ins**, and **Exp**, the local deformation must be accounted in **ExpR**. All **ExpR** voxels with HU values <T were identified on the Jacobian determinant map, **J**, which quantifies the amount of volumetric deformation at the voxel-level (18, 37). QAT and mean HU (mHU) values for a given segmented volume were corrected using **J** based on the following expressions:

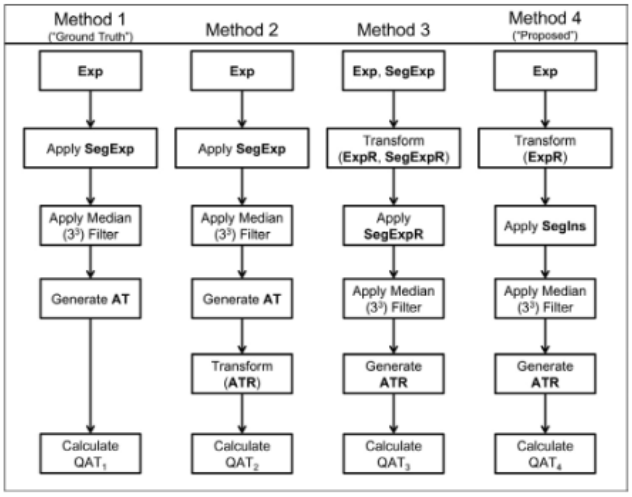

Figure 2. Workflow for calculating QAT using Methods 1–4. Presented are the individual steps for calculating QAT for each method of analysis. Methods 2, 3, and 4 capture effects of local deformation, filtering, and misalignment on QAT, with Method 1 serving as "truth." Method 4 is the "proposed" approach for analysis of spatially aligned quantitative metrics. **Exp** is the original 3D expiration CT scan. **SegExp** and **SegIns** are the lobe segmentation maps generated from the original 3D expiration and inspiration CT scans, respectively. **AT** is the 3D air trapping binary map. QAT_i is the quantitative air trapping scalar value for the ith method. "R" indicates alignment of the imaging data to the inspiration CT frame. Median (3^3) is the 3^3 median filter. The same subject-specific threshold for classifying individual voxels as air trapping was used for all methods. Sixty-four thresholds, 16 subjects with 4 time points, were calculated for this study.

$$\mathrm{QAT(corrected)} = \sum_{i=1}^{N} \left[\mathrm{AT_{ExpR,i}} * \mathrm{J_i}\right] / \sum_{i=1}^{N} \mathrm{J_i} \qquad (4)$$

$$\mathrm{mHU(corrected)} = \sum_{i=1}^{N} \left[\mathrm{ExpR_i} * \mathrm{J_i}\right] / \sum_{i=1}^{N} \mathrm{J_i} \qquad (5)$$

where N is the total number of voxels in the vectorized form of **ExpR** and **J**. QAT values for **Exp** and **ExpR** are presented in this study as percent relative to total segmented volume. This approach is similar to the strategy proposed by Staring and colleagues, which uses the determinant of the Jacobian to account for differences in inflation levels for serial calculation of emphysema (21).

Data and Statistical Analysis

QAT measurements were determined using four different methods of analysis (Fig 2). To differentiate the different QAT, a subscript was included to indicate the method used for the calculation. The first method (Method 1) used the **Exp** with corresponding segmentation map (**SegExp**) to calculate QAT_1. This is the current standard method defined as "ground truth." To demonstrate the effects of parenchymal deformation (Method 2), QAT_2 was calculated using **AT** generated from **Exp** and **SegExp** then transformed to the inspiration space (**ATR**). This method was used to validate our simulated results using experimental data. Effects of imaging filtering (Method 3) were demonstrated by calculating QAT_3

using **ExpR** with corresponding segmentation map (**SegExpR**), which eliminates the effect of misregistration on the measurements. In the fourth method (Method 4) QAT_4 was calculated using the same **ExpR** as in Method 3 but with the segmentation map from the inspiration CT scan (**SegIns**). QAT values determined using this method are susceptible to misalignment of the segmented volumes. Nevertheless, Method 4 is the "proposed" approach for analyzing QAT, as it utilizes the segmentation map from the inspiration CT scan, which is more reliable than from the expiration CT scan (38). For each method, mHU was determined following application of the median filter. Bland-Altman plots were used to assess variability and bias in QAT and HU measurements of Methods 2–4 against Method 1 (i.e. "ground truth") (39). In addition, we evaluated agreement between **SegIns** and **SegExpR** using a Dice coefficient analysis (40). All QAT and mHU values were determined over the lungs and individual lobes: right upper lobe (RUL), right middle lobe (RML), right lower lobe (RLL), left upper lobe (LUL), lingula (LLi) and left lower lobe (LLL). All image processing was performed using in-house algorithms developed in a technical computing language (MATLAB v. R2016a, The MathWorks Inc., Natick, MA). Statistical computations were performed using the Statistics and Machine Learning Toolbox in Matlab (v. R2016a, The MathWorks Inc., Natick, MA).

IRB Approval

The study was approved by the Institutional Review Boards at Stanford University Medical Center and Ohio State University School of Medicine. Full informed consent for the child's assent for examination and further data processing was obtained from the parents or guardians of the children.

RESULTS

Illustration of QAT on Aligned CT

In a single case, we illustrate the use of Method 4 of this study to calculate QAT (i.e. QAT_4 in Fig 2) values for each lobe of lungs (Fig 3). The data used to generate these results were acquired from an 8-year-old female diagnosed with mild CF (FEV1 and FEV1% predicted of 1.76 L and 112%, respectively). At the time of accrual, the subject was found to have QAT of 6.5% as defined using a subject-specific threshold of −673 HU on the unregistered expiration CT scan (**Exp**). By spatially aligning the expiration CT scan to the inspiration geometric frame, we are able to apply the **Ins**-derived lobe segmentation map to the **Exp**-derived **AT** map (**SegIns** and **AT**, respectively, in Fig 3). Inclusion of the airway tree in Figure 3 is to further demonstrate the potential of this "proposed" approach to analyzing multiple quantitative metrics. QAT_4 values varied by lobes (Fig 3 **SegIns AT**) with RML and LUL having the most (14.4%) and least (2.5%) air trapping. These results change when compared to QAT_1 values (i.e. "truth"), where air trapping in RML and LUL were 13.8% and 4.0%, respectively. The potential for error

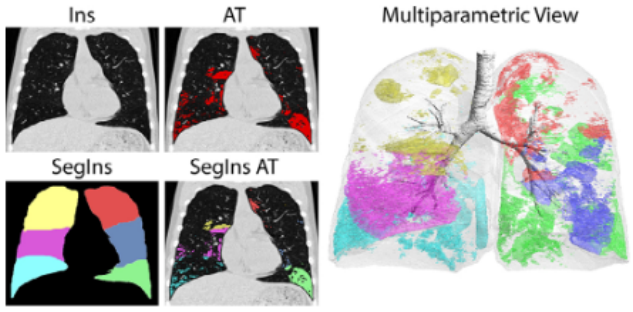

Figure 3. Spatial alignment of QCT metrics to the inspiratory frame (i.e. Method 4). Presented is a representative coronal slice in the inspiratory frame for the CT scan (**Ins**), air trapping map overlaid on **Ins** (**AT**), lobe segmentation map (**SegIns**) and **AT** coded for each lobe overlaid on **Ins** (**SegIns** on **AT**) acquired from a 8-year-old female diagnosed with mild CF at the time of accrual. A multiparametric view of the 3D rendered airway tree and **AT** coded for each lobe allow for visualization and quantification of local disease. Individual lobes presented in the multiparametric view are color-coded as yellow for right upper lobe (RUL), magenta for right middle lobe (RML), cyan for right lower lobe (RLL), red for left upper lobe (LUL), blue for lingula and green for left lower lobe (LLL). (Color version of figure is available online.).

in our QAT measurements may occur as a result of normalizing the sum of all air trapping voxels to the total lung volume at the inspiration frame. In other words, QAT using **Exp** ≠ QAT using **ExpR**.

For the same case, we present in Figure 4 representative images from CT scans acquired at both inflation levels, aligned expiration CT scan and the corresponding determinant of the Jacobian map, which illustrates the extent of local deformation required to spatially align **Exp** to **Ins**. Hypointense regions are clearly evident on **Exp**, a hallmark of air trapping. To spatially align **Exp** to **Ins**, sophisticated registration algorithms are required to address the spatially dependent deformation in the lung parenchyma. Volumes of air trapping, example indicated by arrow in Figure 4, undergo less deformation than what occurs for healthy parenchyma. This data provides some support for the model assumptions used to derive variations in QAT using aligned expiration CT scans (Fig 1).

Effect of Local Deformation on QAT Values: Simulated and Experimental Results

The mathematical expression presented in Equation (3) models the deviation of QAT_{E2I} from QAT_E that would occur when neglecting differences in lung volumes at inspiration and expiration. Presented in Figure 5A are simulated results from Equation (3) for various α. The model predicts no difference in QAT values when there is no air trapping ($QAT_{E2I} = QAT_E = 0\%$), only air trapping ($QAT_{E2I} = QAT_E = 100\%$) and no change in the volume of healthy parenchyma ($\alpha = 1$). For all other cases, QAT_{E2I} was found to underestimate the true value (QAT_E), with the maximum deviation occurring at the average QAT of 50%. In fact, the model predicts that only a 20% change in healthy parenchyma between

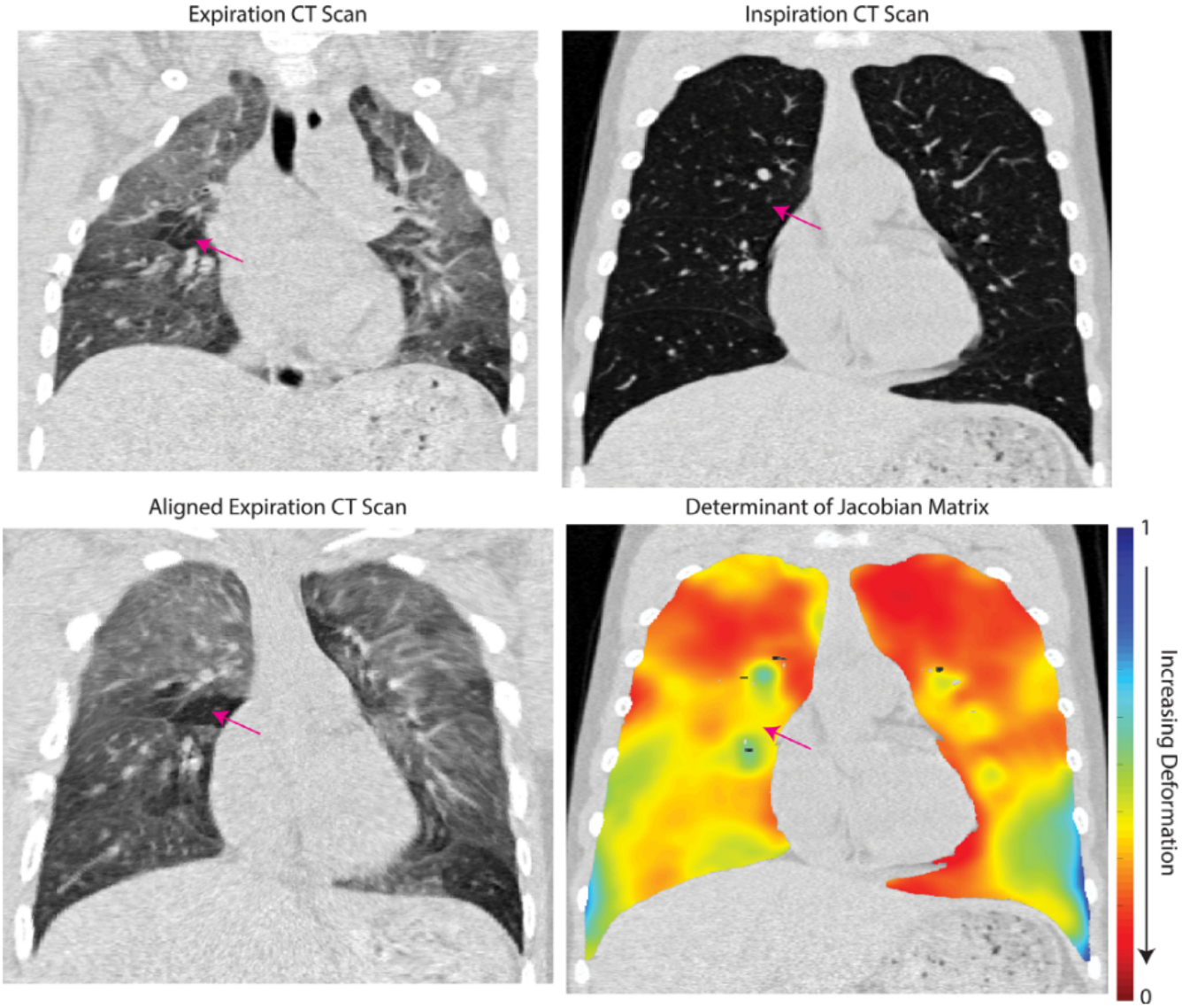

Figure 4. Example of differences in local deformation. Presented are representative coronal slices of the expiration, inspiration, and aligned expiration CT scans with corresponding determinant of the Jacobian matrix map (**J**) from the same subject as in Figure 3. The magenta arrow indicates the location of regional air trapping on all image slices. Spatial alignment of air-trapped regions requires less deformation than hyperattenuated parenchyma as indicated by dark red regions on the J map. For reference, J = 1 for no deformation, decreasing values ($0 < \mathbf{J} < 1$) for expansion and increasing values ($\mathbf{J} > 1$) for compression. (Color version of figure is available online.).

inflation levels ($\alpha = 0.8$) is required to result in a difference in QAT values of 5%.

Next, we demonstrated the deviation in QAT calculated from aligned expiration CT scans using our clinical data. This was performed by comparing whole-lung QAT values calculated using Methods 1 (QAT_1) and 2 (QAT_2) for all cases and time points. Figure 5B (orange markers) shows results with similar trends to those observed in our simulated data (Fig 5A). As predicted by the model, QAT_2 underestimated QAT_1, which was found to differ by as much as 5%. Our experimental results did deviate from what we would have predicted from our simple model. First, the maximum deviation in simulated QAT values occurred at an average QAT of 50%, whereas our experimental results peaked between 30% and 40%. As observed in Figure 4, local deformation was evident in regions of air trapping. In fact, the lung volume at expiration was, on average ± the standard deviation, 0.36 ± 0.1 of the lung volume at inspiration. Nevertheless, the deviation in QAT values may only be attributed to the deformation needed to align **Exp** to **Ins**, and as such, can be correct using the **J**. Calculation of QAT using the **J** (Eq. 4) removed nearly all deviation in QAT values (Fig 5B, blue markers).

Effects of Filtering and Alignment on QAT Values

We sought to determine if other confounding factors alter QAT values following image registration of **Exp** to **Ins** space. Figure 6 shows Bland-Altman plots used to evaluate QAT and

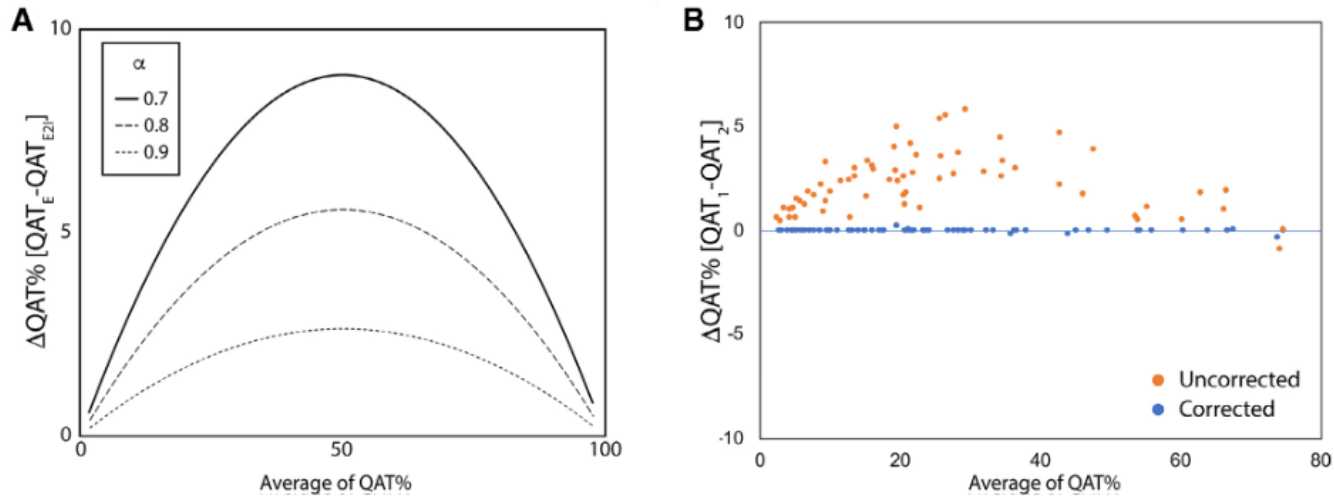

Figure 5. Local deformation alters QAT values. Presented are simulated and experimental results demonstrating the effect of local deformation on calculations of QAT using the common definition of volume of air trapping normalized by total segmented volume. Bland-Altman plots are presented to compare QAT values using (A) the simulated QAT_{E2I}, as expressed in Equation (3), and QAT_E ("truth") for varying values of α, the ratio of healthy parenchyma volume at expiration and inspiration and (B) whole-lung QAT_2 and QAT_1 ("truth") as described in Figure 2. The model QAT variables with indices E and E2I indicate values determined from the unregistered and registered expiration CT scan, respectively (Fig 1). Orange and blue markers indicate QAT values that are uncorrected and corrected, respectively, for local deformation using **J** (Eq. 4). (Color version of figure is available online.).

mHU values determined from the registered expiration CT scan (**ExpR**) and the unaligned scan (**Exp**). Two separate analyses were performed. The first, presented in Figure 6A and 6B, compares QAT_3 and mHU_3 to QAT_1 and mHU_1, respectively. This approach removes the effect of misalignment on our analysis, leaving only image filtering as a possible confounder of measurement variability. Shown in Figure 6A, a positive bias in the uncorrected QAT_3 was observed, consistent with the simulations and experimental results presented in Figure 5. In contrast, variability in the uncorrected mHU values was observed with a slight negative bias for increasing values (Fig 6B). The variability in mHU values was found to increase for values greater than the maximum subject-specific threshold determined from all paired CT scans ($N = 64$)(Fig 6B small dashed line). As expected, QAT_3 and mHU_3 corrected using J (Eqs. 4 and 5) eliminated most of the variability observed in both measures. Presented in Table 2 is a tabulation of cases that demonstrate deviations greater than 2.5% and 5% in uncorrected and corrected QAT evaluated for the whole-lung and each lobe. In general, applying the **J** removed nearly all deviations > 2.5% from QAT_3. This suggests that application of a 3^3 median filter to **ExpR** has a negligible effect on the variability of the corrected QAT.

The contribution of registration misalignment on the variability of volume-based QAT and mHU measurements was determined by comparing the measurements from Method 4 ("proposed") to those of Method 1 ("truth"). As seen in Figure 6C, the uncorrected and corrected QAT measurements were found to have more scatter than what was observed for filtering effects alone (Fig 6A). Uncorrected QAT_4, calculated over the whole-lung, with deviations > ±2.5% and > ±5% were observed in 29 and 5 cases, respectively (Table 2). Correcting for inflation level volumes eliminated all deviations >2.5%. Performing this same analysis over the lobes, we found the highest variability in the corrected QAT_4 in the right middle lobe (11 out of 64 were >2.5%) and the lingula (21 out of 64 were >2.5%). Contrary to QAT_4, correcting for lung volume at each inflation level was found to have little benefit in mitigating the bias and deviations in whole-lung mHU_4 values (Fig 6D), with deviations lower than −60HU for average values > −500HU. Although there was significant variability in mHU_4 from those of mHU_1, most of the variability occurred at HU values greater than the upper limit threshold of −570HU (Fig 6D small dashed line).

Agreement Between Aligned Expiration and Inspiration Segmentation Maps

As both QAT_4 and mHU_4 were calculated over segmented volumes (i.e. whole lung or individual lobes), one would not suspect misalignment at the voxel-level to have a profound effect on these values. Instead, inconsistencies in the segmented volumes at inspiration and expiration, typically along the border of the segmented volumes, would result in variability in the measurements. To determine the extent of misalignment at the whole-lung and lobe-level, we performed a Dice Coefficient analysis to assess agreement between the segmented volumes determined from **Ins** and **ExpR** (**SegIns** and **SegExpR**, respectively). As seen in Figure 7A, clear differences are evident along the border of the segmented lobes for the same case used in Figures 3 and 4. The most striking is the border between the lingula and left lower lobe (blue and green segments, respectively, in Fig 7A), which is attributed

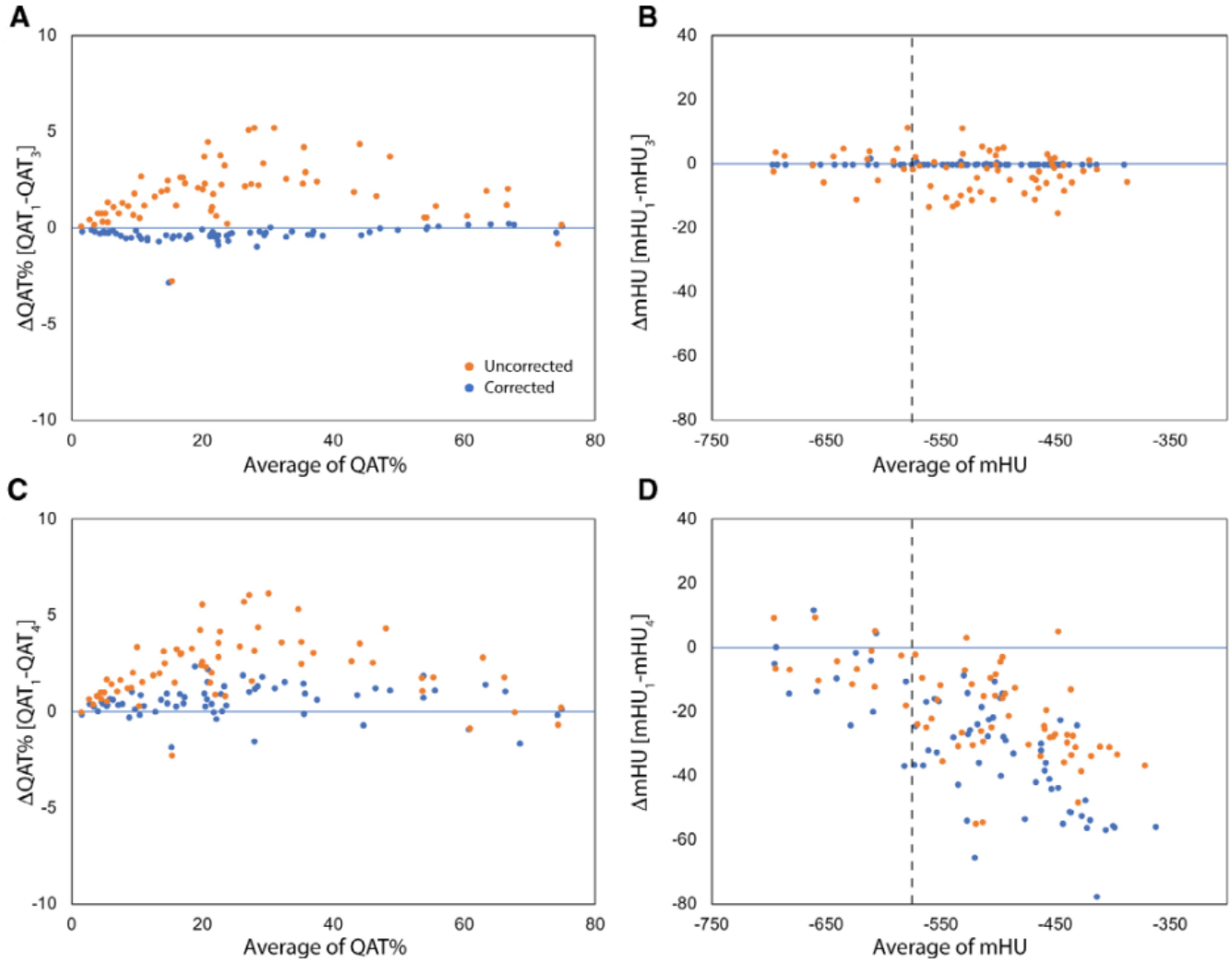

Figure 6. Effect of image registration on quantitative measurements. Presented are comparisons of whole-lung QAT and mHU values generated using Methods 3 and 4 to like values from Method 1 (i.e. "ground truth"). Bland-Altman plots are presented to compare metric values using (A and B) Method 3 to Method 1 and (C and D) Method 4 to Method 1. Orange and blue markers indicate QAT and mHU determined using Methods 3 and 4 uncorrected and corrected, respectively, for local deformation using **J** (Eqs. 4 and 5, respectively). The vertical dashed line indicates the highest subject-specific threshold calculated for classifying regions of parenchyma as air trapping. This threshold was identified as −570 HU. (Color version of figure is available online.).

TABLE 2. Number of Cases With Variability That Exceed 2.5% and 5%

Segment	$\lvert QAT_1 - QAT_3 \rvert$				$\lvert QAT_1 - QAT_4 \rvert$			
	>2.5%		>5%		>2.5%		>5%	
	Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected
RUL	0	15	0	0	9	25	1	8
RML	0	1	0	1	11	13	1	2
RLL	0	15	0	2	0	9	0	2
LUL	0	10	0	1	5	16	1	3
LLi	1	2	1	1	21	15	10	8
LLL	0	7	0	2	2	9	0	2
Whoe Lung	0	16	0	0	0	29	0	5

Note: Data is presented as cases where the absolute difference in QAT_3 and QAT_4 values from QAT_1 ("truth") exceeded set limits. All 64 cases were evaluated for whole-lung and individual lobes. Corrected and Uncorrected indicate whether the determinant of the Jacobian was used to correct for local deformation or not, respectively. Individual lobes are right upper lobe (RUL), right middle lobe (RML) and right lower lobe (RLL); left upper lobe (LUL), lingula (LLi), and lower left lobe (LLL).

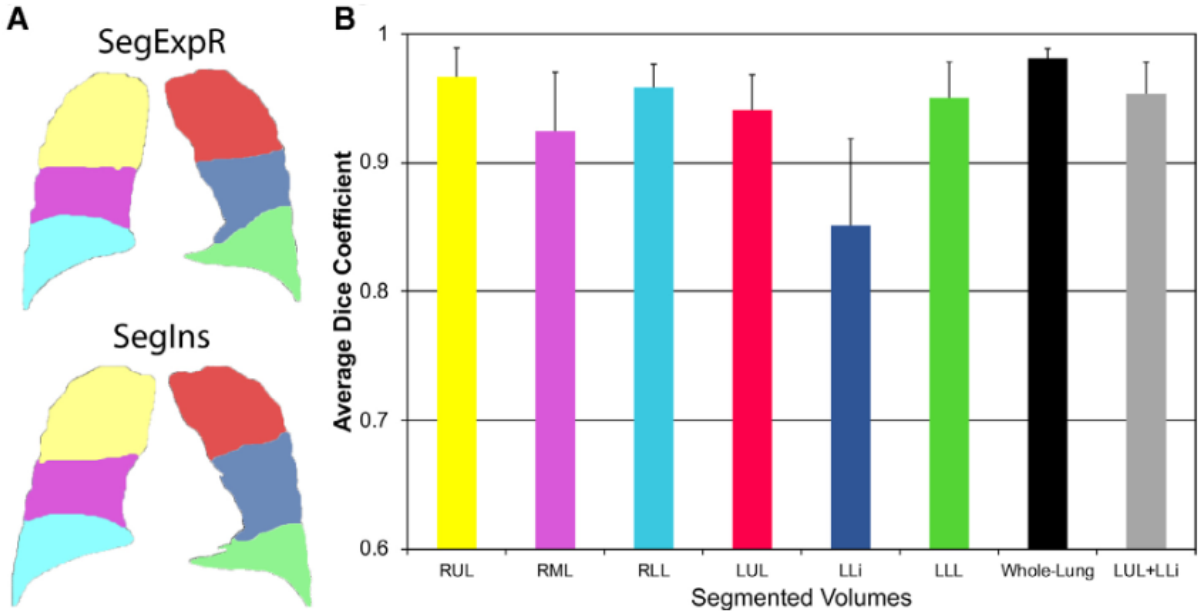

Figure 7. Agreement between aligned expiration and inspiration lobe segmentation maps. Dice Coefficient analysis was performed to determine the agreement between segmentation maps produced on the expiration CT scan transformed to the inspiration space [(A Top); **SegExpR**] and the inspiration CT scan [(A Bottom); **SegIns**]. Representative coronal slice of the segmented maps are presented from the same subject as Figures 3 and 4. (B) Bar plot shows all results for all paired CT ($N = 64$) for each lobe and for the whole-lung. Data is presented as mean ± standard deviation. Individual lobes presented in (A) and bar plot presented in (B) are color-coded as yellow for right upper lobe (RUL), magenta for right middle lobe (RML), cyan for right lower lobe (RLL), red for left upper lobe (LUL), blue for lingula (LLi), green for left lower lobe (LLL), black for whole-lung and gray for the union of LUL and LLi segmented volumes. (Color version of figure is available online.).

to large deformations in the inferior portions of the lungs. Nevertheless, large volumes of the lobes were found to be in good agreement when defined by **SegIns** and **SegExpR** (Fig 7B). The mean Dice Coefficient (DC) was found to be >0.9 for nearly all segmented volumes. Only the lingula was found to have a mean DC lower (mean DC = 0.85 ± 0.07), yet still considered in good agreement between **SegIns** and **SegExpR**. The lingula is typically considered part of the left upper lobe. Merging the segmented volume of the lingula with the left upper lobe resulted in a DC of 0.95 ± 0.02, which is consistent with the other lobes.

DISCUSSION

Extensive research has been devoted to developing quantitative CT metrics for accurate diagnostics and prognostication of lung diseases. At present, a variety of QCT approaches has been developed but are restricted to the geometric frame of the CT scan from which they were derived. Individually these QCT metrics provide functional and spatial information that has been shown to be efficacious. Examples include the local distribution of emphysema in the inferior lung a characteristic of $\alpha-1$ antitrypsin deficiency (41), feature detection of idiopathic pulmonary fibrosis (42), and topology of quantitative air trapping as a measure of small airways disease severity (43). Nevertheless, analysis of multiple QCT metrics from different scans typically involves simplification of the 3D maps into scalar quantities allowing easy interpretation and comparison between metrics. As a consequence, the spatial information from which these scalar quantities were derived is lost. Image registration of serial CT scans has overcome this limitation, allowing investigators to fully exploit the spatial-dependent details within these 3D maps. Although analytical techniques that implement registration are becoming more widely available, little is known as to the effect this post-processing step has on QCT metrics, such as quantitative air trapping. The motivation of this work was to assess sources of variability on QAT measurements calculated from registered expiration CT scans. Using CT data from a well-controlled cohort of CF patients and visually inspected lobe segmentation maps, we were able to evaluate the effect of image registration on quantitative CT measurements. We observed that significant deviation in QAT values may occur but is resolved using a simple correction approach.

Contribution of Local Deformation on QAT

The most significant variability in QAT values was found to be attributed to neglecting parenchymal volume differences between inflation levels (Fig 5). These observations are due to differences in the extent of deformation between the air trapped and healthy regions of the lung. As illustrated in Figure 4, lung deformation varies spatially and is dependent on the presence of air trapping. Dense regions on CT scans at expiration are characteristic of normally ventilated lung parenchyma that may more fully deflate on exhalation (7). The red regions on the Jacobian determinant map ($\mathbf{J}$) in Figure 4 require significant deformation of the expiration CT

scan to spatially align to the fully-inflated lungs captured on CT scans at inspiration (17, 18). In our simplified mathematical model (Fig 1 and Eq. 3), we were able to simulate the variability in QAT_{E2I} from QAT_E due to the differences in volume changes for healthy and air-trapped parenchyma (Fig. 1 and 5A). Variability in the clinical QAT measurement was found to peak at lower mean QAT values than what was predicted by the mathematical model (Fig 5). The discrepancy between simulated and experimental results is most likely attributed to the model assumption that air trapping volumes are incompressible (no difference in volume at inspiration and expiration; $\mathbf{J} = 1$). Nevertheless, our simulations were corroborated by our clinical QAT_2 results (Fig 5B). The extent of lung deformation, as measured by **J**, has been found to correlate strongly with pulmonary function. Reinhardt and colleagues showed in an animal model that increased lung pressure (i.e. inflation) was linearly proportional to lung deformation (18). Although a bias of up to 5% was observed in QAT (Fig. 5B, 6A and C, Table 2), most of this variability was removed when using **J** in Equation 4 to calculate QAT from the aligned expiration CT scan. It is important to note that the effect of local deformation on QAT values should be independent of patient characteristics or extent of disease. The variability in QAT using registered CT data is simply associated with the difference in inflation level between the target, here **Ins**, and moving, here **Exp**, CT data sets. The higher the difference in inflation levels between **Ins** and **Exp** the more deformation required to spatially align the data, resulting in more variability in QAT measured using aligned CT data to the original CT data.

Effect of Misalignment on QAT

Misalignment of **Exp**-derived to **Ins**-derived lobe segmentation maps was found to be the second highest contributor of variability in the evaluated metrics. In the same study by Reinhardt and colleagues, they observed increased misregistration with increasing inflation (18). This observation may explain the increased scatter observed in our QAT and bias in mHU (Fig 6C and D). As discussed previously, regions of high attenuation on expiration CT scans (Fig 4) are due to deflated parenchyma, which require significant deformation to align with the fully-inflated lungs on the inspiration CT scans. In this study, we observed that elevated mHU measurements, highly attenuated, result in a negative bias of up to 60HU between values calculated using **ExpR** (registered) and **Exp** (original) (Fig 6D). Correcting local deformation had little effect on the difference in mHU values. As QAT and mHU were calculated over segmented volumes, the effect of misregistration on these metrics occurred primarily due to misalignment of the segment borders between **SegIns** and **SegExpR** (Fig 7A). This explains the low values in Dice Coefficients for the RML and, to a greater extent, lingula, as they are both situated within the center of each lung. The difficulty in accurately aligning these lobes is evident in the large number of cases where corrected QAT_4 exceeded ±2.5% (Table 2). This effect will diminish as new, more accurate registration algorithms become available for QCT analysis.

Contribution of Threshold on QAT from Aligned Expiration CT

The QAT metric is dependent on the threshold HU value used to identify which voxels are air trapping in the expiration CT scan. When calculating QAT, large threshold values are suspected of including regions of parenchyma that may undergo significant volumetric change from inspiration to expiration. Our study cohort was from a young population of CF patients, where all CT scans were spirometrically-gated to total lung capacity and near residual volume, which is a deeper expiration than normally achieved in expiratory imaging at functional residual capacity or in routine imaging without gating, and therefore a higher overall CT lung density is achieved on expiration. For QAT, we employed a subject-specific thresholding approach for the detection of mild-to-severe CF, a method developed by Goris and colleagues (8). This approach was selected over other common air trapping thresholding values, such as -856 HU for COPD (44), as it has been shown to provide a better estimation of the extent of QAT for this young CF patient population (8). We calculated thresholds in the range of −738 HU to −570 HU (upper limit indicated by the vertical small dashed line in Fig 6B and D) with a median of −676 HU. Although high (more dense) compared to the COPD accepted threshold of −856 HU, deviations of mHU_4 from mHU_1 were approximately ±20HU for average values less than −570HU (vertical line in Fig 6D), which indicate regions of air trapping. This may account for the negligible effect misregistration had on QAT_4 values (Fig 6C). Nevertheless, care must still be taken when interpreting CT density values from registered CT data.

LIMITATIONS

It is important to address some of the limitations of this study. Typical approaches for determining registration accuracy are through comparison of landmarks between the moving and target images (15, 29) (**Exp** and **Ins**, respectively). Though this is a more precise approach for measuring misregistration, this analysis was not required as our study was concerned with variability in QAT and mean HU values calculated over segmented volumes. An important attribute of this cohort was the availability of manually corrected lobe segmentation maps for all paired CT scans. By focusing our attention to the segmentation maps, we were able to explore the potential of lobe-based inspiratory segmentation maps for expiratory QAT calculations. Algorithms developed to segment the lobes on CT scans acquired at inspiration are currently being investigated for lobe segmentation on expiration CT scans (14). While promising, a fully automated lobe segmentation mapping technique is complicated by the large dynamic range in parenchymal densities observed in

expiration CT scans resulting in less reliable results (38). To the best of our knowledge, we have demonstrated for the first time the application of image registration for QAT lobe analysis using segmentation maps derived from inspiration CT scans (Method 4, Fig 3). By manually inspecting each lobe segmentation map we have "ground truth" in our QAT calculations at the whole-lung and lobe levels. Overall, this study demonstrates that a bias in QAT values from aligned CT data can be mitigated by using a simple approach. In view of inter-reader variability or subjectivity of visual semi-quantification of air trapping by current scoring systems (45–48), we find these results highly valuable in that an error rate of roughly 5% for a lobe-based quantification of air trapping may be acceptable for future trials, and also in clinical routine application.

CONCLUSIONS

We present here our results on evaluating sources of variability in QAT values that may occur due to the spatial alignment of expiration CT scan to the inspiration CT geometric frame. Accomplishment of this analysis was achieved through the use of paired CT data, with corresponding visually inspected lobe segmentation maps, from a well-controlled cohort of CF subjects. Even in the presence of large deviation in HU values key sources of variability in image registration only resulted in mild differences in QAT values when compared to "truth," well within scientifically and clinically acceptable margins. Based on our findings, the most significant source of variability was when local deformation was neglected when calculating QAT. Simply applying the determinant of the Jacobian matrix in calculating QAT resolved this issue. Novel analytical techniques that incorporate image registration are providing unique information about obstructive lung diseases. These spatially-resolved multiparametric methods may soon integrate vascular, parenchymal and airway changes in lung disease on an identical geometric frame. This will allow researchers and clinicians to further exploit the wealth of functional and anatomical information that resides in the various QCT metrics, resulting in improved clinical-care of patients suffering from lung diseases.

REFERENCES

1. Herth FJF, Kirby M, Sieren J, et al. The modern art of reading computed tomography images of the lungs: quantitative CT. Respiration 2018; 95 (1):8–17.
2. Lynch DA, Newell JD. Quantitative imaging of COPD. J Thorac Imaging 2009; 24:189–194.
3. Wielputz MO, Weinheimer O, Eichinger M, et al. Pulmonary emphysema in cystic fibrosis detected by densitometry on chest multidetector computed tomography. PLoS One 2013; 8:e73142.
4. Mascalchi M, Camiciottoli G, Diciotti S. Lung densitometry: why, how and when. J Thorac Dis 2017; 9:3319–3345.
5. Coxson HO, Rogers RM, Whittall KP, et al. A quantification of the lung surface area in emphysema using computed tomography. Am J Respir Crit Care Med 1999; 159:851–856.
6. Gevenois PA, De Vuyst P, de Maertelaer V, et al. Comparison of computed density and microscopic morphometry in pulmonary emphysema. Am J Respir Crit Care Med 1996; 154:187–192.
7. Newman KB, Lynch DA, Newman LS, et al. Quantitative computed tomography detects air trapping due to asthma. Chest 1994; 106:105–109.
8. Goris ML, Zhu HJ, Blankenberg F, et al. An automated approach to quantitative air trapping measurements in mild cystic fibrosis. Chest 2003; 123:1655–1663.
9. Matsuoka S, Kurihara Y, Yagihashi K, et al. Quantitative assessment of air trapping in chronic obstructive pulmonary disease using inspiratory and expiratory volumetric MDCT. AJR Am J Roentgenol 2008; 190:762–769.
10. Hammond E, Sloan C, Newell Jr. JD, et al. Comparison of low- and ultra-low-dose computed tomography protocols for quantitative lung and airway assessment. Med Phys 2017; 44:4747–4757.
11. Lim HJ, Weinheimer O, Wielputz MO, et al. Fully automated pulmonary lobar segmentation: influence of different prototype software programs onto quantitative evaluation of chronic obstructive lung disease. PLoS One 2016; 11:e0151498.
12. Leutz-Schmidt P, Weinheimer O, Jobst BJ, et al. Influence of exposure parameters and iterative reconstruction on automatic airway segmentation and analysis on MDCT-An ex vivo phantom study. PLoS One 2017; 12:e0182268.
13. Weinheimer O, Wielpütz MO, Konietzke P, et al. Fully automated lobe-based airway taper index calculation in a low dose MDCT CF study over 4 timepoints. Medical Imaging 2017: Image Processing. Orlando, Florida, United States: SPIE, 2017.
14. Gerard SE, Patton TJ, Christensen GE, et al. FissureNet: a deep learning approach for pulmonary fissure detection in CT images. IEEE Trans Med Imaging 2018. doi:10.1109/TMI.2018.2858202. [Epub ahead of print].
15. Cunliffe AR, White B, Justusson J, et al. Comparison of two deformable registration algorithms in the presence of radiologic change between serial lung CT scans. J Digit Imaging 2015; 28:755–760.
16. Shamonin DP, Bron EE, Lelieveldt BP, et al. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. Front Neuroinform 2013; 7:50.
17. Bhatt SP, Bodduluri S, Hoffman EA, et al. Computed tomography measure of lung at risk and lung function decline in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2017; 196:569–576.
18. Reinhardt JM, Ding K, Cao K, et al. Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation. Med Image Anal 2008; 12:752–763.
19. Nishio M, Matsumoto S, Tsubakimoto M, et al. Paired inspiratory/expiratory volumetric CT and deformable image registration for quantitative and qualitative evaluation of airflow limitation in smokers with or without COPD. Acad Radiol 2015; 22:330–336.
20. Lee SM, Seo JB, Lee SM, et al. Optimal threshold of subtraction method for quantification of air-trapping on coregistered CT in COPD patients. Eur Radiol 2016; 26:2184–2192.
21. Staring M, Bakker ME, Stolk J, et al. Towards local progression estimation of pulmonary emphysema using CT. Med Phys 2014; 41:021905.
22. Gorbunova V, Jacobs SS, Lo P, et al. Early detection of emphysema progression. Med Image Comput Comput Assist Interv 2010; 13:193–200.
23. Belloli EA, Degtiar I, Wang X, et al. Parametric response mapping as an imaging biomarker in lung transplant recipients. Am J Respir Crit Care Med 2017; 195:942–952.
24. Galban CJ, Boes JL, Bule M, et al. Parametric response mapping as an indicator of bronchiolitis obliterans syndrome after hematopoietic stem cell transplantation. Biol Blood Marrow Transplant 2014; 20:1592–1598.
25. Galban CJ, Han MK, Boes JL, et al. Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. Nat Med 2012; 18:1711–1715.
26. Pompe E, Galban CJ, Ross BD, et al. Parametric response mapping on chest computed tomography associates with clinical and functional parameters in chronic obstructive pulmonary disease. Respir Med 2017; 123:48–55.
27. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. COPD 2010; 7:32–43.
28. Sieren JP, Newell Jr. JD, Barr RG, et al. SPIROMICS protocol for multicenter quantitative computed tomography to phenotype the lungs. Am J Respir Crit Care Med 2016; 194:794–806.
29. Murphy K, van Ginneken B, Reinhardt JM, et al. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. IEEE Trans Med Imaging 2011; 30:1901–1920.

30. Robinson TE, Goris ML, Moss RB, et al. EPS4.3 Differential sensitivity of outcome measures that assess progression of mild CF lung disease in school age children. J Cystic Fibrosis 2017: S44–S45.
31. Wielputz MO, Eichinger M, Weinheimer O, et al. Automatic airway analysis on multidetector computed tomography in cystic fibrosis: correlation with pulmonary function testing. J Thorac Imaging 2013; 28:104–113.
32. Weinheimer O, Achenbach T, Bletz C, et al. About objective 3-d analysis of airway geometry in computerized tomography. IEEE Trans Med Imaging 2008; 27:64–74.
33. Health Risks from Exposure to Low Levels of Ionizing Radiation. National Academies Press, 2006.
34. Kuo W, Ciet P, Tiddens HAWM, et al. Monitoring cystic fibrosis lung disease by computed tomography. Radiation risk in perspective. Am J Respir Crit Care Med 2014; 189:1328–1336.
35. Klein S, Staring M, Murphy K, et al. Elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010; 29:196–205.
36. Hatt CR, Galban CJ, Humphries SM, et al. Noise Reduction Filtering Improves Repeatability of Quantitative Lung Density Metrics Between Full and Reduced-dose CT Scans. Chicago, IL: RSNA, 2016.
37. Reinhardt JM, Christensen GE, Hoffman EA, et al. Registration-derived estimates of local lung expansion as surrogates for regional ventilation. Inf Process Med Imaging 2007; 20:763–774.
38. Konietzke P, Weinheimer O, Wielputz MO, et al. Validation of automated lobe segmentation on paired inspiratory-expiratory chest CT in 8-14 year-old children with cystic fibrosis. PLoS One 2018; 13:e0194557.
39. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015; 25:141–151.
40. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging 2015; 15:29.
41. Parr DG, Stoel BC, Stolk J, et al. Pattern of emphysema distribution in alpha1-antitrypsin deficiency influences lung function impairment. Am J Respir Crit Care Med 2004; 170:1172–1178.
42. Jacob J, Bartholmai BJ, Rajagopalan S, et al. Automated quantitative computed tomography versus visual computed tomography scoring in idiopathic pulmonary fibrosis: validation against pulmonary function. J Thorac Imaging 2016; 31:304–311.
43. Hoff BA, Pompe E, Galban S, et al. CT-Based local distribution metric improves characterization of COPD. Sci Rep 2017; 7:2999.
44. Nambu A, Zach J, Schroeder J, et al. Quantitative computed tomography measurements to evaluate airway disease in chronic obstructive pulmonary disease: relationship to physiological measurements, clinical index and visual assessment of airway disease. Eur J Radiol 2016; 85: 2144–2151.
45. Rosenow T, Oudraad MC, Murray CP, et al. PRAGMA-CF. A quantitative structural lung disease computed tomography outcome in young children with cystic fibrosis. Am J Respir Crit Care Med 2015; 191:1158–1165.
46. Loeve M, Gerbrands K, Hop WC, et al. Bronchiectasis and pulmonary exacerbations in children and young adults with cystic fibrosis. Chest 2011; 140:178–185.
47. Robinson TE, Leung AN, C`hen X, et al. Cystic fibrosis HRCT scores correlate strongly with Pseudomonas infection. Pediatr Pulmonol 2009; 44:1107–1117.
48. Brody AS, Kosorok MR, Li Z, et al. Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. J Thorac Imaging 2006; 21:14–21.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found at https://doi.org/10.1016/j.acra.2018.11.001.