Diagn Interv Radiol 2013; 19:279–285

CHEST IMAGING
ORIGINAL ARTICLE

Intra- and interoperator variability of lobar pulmonary volumes and emphysema scores in patients with chronic obstructive pulmonary disease and emphysema: comparison of manual and semi-automated segmentation techniques

Francesco Molinari, Tommaso Pirronti, Nicola Sverzellati, Stefano Diciotti, Michele Amato, Guglielmo Paolantonio, Luigia Gentile, George K. Parapatt, Francesco D'Argento, Jan-Martin Kuhnigk

PURPOSE
We aimed to compare the intra- and interoperator variability of lobar volumetry and emphysema scores obtained by semi-automated and manual segmentation techniques in lung emphysema patients.

MATERIALS AND METHODS
In two sessions held three months apart, two operators performed lobar volumetry of unenhanced chest computed tomography examinations of 47 consecutive patients with chronic obstructive pulmonary disease and lung emphysema. Both operators used the manual and semi-automated segmentation techniques. The intra- and interoperator variability of the volumes and emphysema scores obtained by semi-automated segmentation was compared with the variability obtained by manual segmentation of the five pulmonary lobes.

RESULTS
The intra- and interoperator variability of the lobar volumes decreased when using semi-automated lobe segmentation (coefficients of repeatability for the first operator: right upper lobe, 147 vs. 96.3; right middle lobe, 137.7 vs. 73.4; right lower lobe, 89.2 vs. 42.4; left upper lobe, 262.2 vs. 54.8; and left lower lobe, 260.5 vs. 56.5; coefficients of repeatability for the second operator: right upper lobe, 61.4 vs. 48.1; right middle lobe, 56 vs. 46.4; right lower lobe, 26.9 vs. 16.7; left upper lobe, 61.4 vs. 27; and left lower lobe, 63.6 vs. 27.5; coefficients of reproducibility in the interoperator analysis: right upper lobe, 191.3 vs. 102.9; right middle lobe, 219.8 vs. 126.5; right lower lobe, 122.6 vs. 90.1; left upper lobe, 166.9 vs. 68.7; and left lower lobe, 168.7 vs. 71.6). The coefficients of repeatability and reproducibility of emphysema scores also decreased when using semi-automated segmentation and had ranges that varied depending on the target lobe and selected threshold of emphysema.

CONCLUSION
Semi-automated segmentation reduces the intra- and interoperator variability of lobar volumetry and provides a more objective tool than manual technique for quantifying lung volumes and severity of emphysema.

From the Department of Radiology (F.M. ✉ *francescomolinari.dr@gmail.com*), C. H. R. U. Lille, France; the Department of Bioimaging and Radiological Sciences (T.P., M.A., G.K.P., F.D.), Catholic University of Rome, Rome, Italy; the Department of Clinical Sciences (N.S.), Division of Radiology, University of Parma, Parma, Italy; the Department of Clinical Physiopathology (S.D.), Computational Biomedical Imaging Laboratory Radiodiagnostic Section, University of Florence, Florence, Italy; the Vascular and Interventional Radiology Unit (G.P.), Pediatric Hospital Bambino Gesù, Rome, Italy; the Department of Radiology (L.G.), Hospital San Maurizio, Bozen, Italy; MeVis Center for Medical Diagnostic Systems and Visualization (J.K.), Bremen, Germany.

Received 7 September 2012; revision requested 10 October 2012; revision received 25 November 2012; accepted 27 November 2012.

Published online 15 February 2013
DOI 10.5152/dir.2013.047

Chronic obstructive pulmonary disease (COPD) and emphysema are important causes of morbidity and mortality in smokers (1). In patients with COPD and extensive emphysema, the lungs globally overinflate and permanently lose their elastic recoil, which contributes to limiting expiratory airflow. In a subgroup of patients who present heterogeneously distributed emphysema at computed tomography (CT), surgical or bronchoscopic treatments may reduce the volume of overinflated lungs, thereby improving the ventilation of the adjacent non emphysematous regions (2, 3).

Quantitative CT methods that measure lobar volumes and density play a key role in defining the regional distribution and severity of emphysema and in determining the target lobe to treat with lung volume reduction (2, 3). CT lobar volumetry also provides a noninvasive method for monitoring the effects of volume-reduction treatments and assessing the functional improvement of the non emphysematous lung (4). In both the preoperative assessment and postoperative monitoring of patients with severe emphysema, the reliability of CT lobar volumetry and emphysema scoring is crucial (2–6). Quantitative CT techniques for emphysema scoring are objective and therefore preferable to visual analysis (7). Recent studies have also demonstrated that CT lobar volumetry has good inter scan reproducibility (8–10). Among the different sources of variability of CT lung volumetry and emphysema scoring, those depending on the method used for lobar segmentation have received little attention (5).

Segmentation is defined as the procedure of partitioning an image into constituent regions (11). To measure the lobar volumes, an image-based segmentation procedure preliminarily identifies the pulmonary lobes (12). The anatomic landmark between two lobes is the pulmonary fissure. Unfortunately, the fissure may have different shapes and variants, be anatomically incomplete, and overall appear ill-defined on CT (13). When the interlobar boundary is unclear, the segmentation of the pulmonary lobes becomes cumbersome (12). Potential segmentation errors at the interlobar boundary shift parenchymal volume between the lobes. If the segmentation error extends over the entire interlobar boundary, the miscategorized volume can alter the proportions among the measured lobar volumes. Repeated segmentations also indicate that the variability of the measured lobar volumes increases when the interlobar landmark is uncertain (10). Particularly in diseased lungs, where pathologic changes deform the fissures, the accuracy of the segmentation is critical for correctly measuring the lobar volumes (12).

The segmentation technique that provides the standard of reference for the lobar volumetry, for validating other emerging segmentation methods, uses manually traced contours to separate the pulmonary

lobes (11, 14, 15). The manual technique allows to correct segmentation errors interactively and revise the volumetric models of the lobes until they are properly defined. However, when applied recursively to a whole chest multidetector CT (MDCT) dataset containing hundreds of images, the manual segmentation is time-consuming and overall clinically impractical (11). In this setting, the manual technique is subject to substantial error sources, which yield variability in the lobar volumes (10). Conversely, computer-based techniques that use mathematical algorithms such as thresholding, pattern recognition, and deformable models facilitate the segmentation (11). The automated segmentation, however, does not fully adapt to the various anatomic shapes and pathologic changes of the lobes, which may result in inconsistent lobar volumetry (11). A semi-automated technique that combines the advantages of computer-aided segmentation and the possibility of manual refinements has recently emerged as an alternative to both techniques (12). The clinical feasibility of semi-automated lobar segmentation has been demonstrated in a large patient series (16).

The variability of the lobar volumes derived from the semi-automated segmentation technique is currently unknown. Moreover, it is not clear whether in pathologic lungs, such as in emphysema patients, a semi-automated or manual segmentation technique should be used to ensure precise measurements of the lobar volumes and CT densities. We hypothesized that the semi-automated segmentation might reduce the intra- and interoperator variability of the lobar volumetry.

Therefore, the objective of our study was to evaluate the intra- and interoperator variability of lobar volumetry obtained by semi-automated segmentation with the variability of volumetry obtained using the manual approach in patients with lung emphysema.

Materials and methods

This retrospective study was performed in 2009 after receiving approval from the institutional review board. Informed consent was waived. Forty-seven unselected consecutive patients who had been classified according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria into stages I–IV and had undergone unenhanced chest CT for clinical evaluation were included in the study (Table 1). Two thoracic radiologists with ten years of experience in chest imaging (F.M. and N.S.) reviewed the 47 CT examinations of this study cohort. All of the CT examinations (LightSpeed Pro 16, GE Healthcare, Milwaukee, Wisconsin, USA) were obtained with a standardized imaging protocol that included a full-inspiration breath-hold helical scan from the apex to base of the lungs with 16×1.25 mm collimation, 1.375:1 pitch, 120 kVp, modulated dose according to the size and anatomic shape of the patient, 0.5 s rotation time, 40–44 cm FOV, 512×512 matrix size, and reconstruction of axial images at 1.25/1.25 mm of thickness/interval with a high spatial-frequency algorithm. All of the chest CT scans were preliminarily evaluated for the presence of lobar anatomy variants.

Table 1. Patients' characteristics and GOLD stages

	GOLD stages				
	Mild (I) n=5	Moderate (II) n=15	Severe (III) n=15	Very severe (IV) n=12	Overall n=47
Gender (Male/Female)	3/2	10/5	12/3	9/3	34/13
Age (years)	62±7.5	69±17	68.9±15	71.6±15.5	68.9±18.5
BMI (kg/m^2)	26.5±6	26.9±9.35	26.3±5	21.3±4.4	25.1±9.85
Smoking (Pk/y)	17.5±2.5	50.8±30	59.9±70	55.9±37	53.8±72.5
FEV1 (%)	97±8.5	64.1±14	38.9±9.5	22.4±6	46.7±45.5
FVC (%)	102±14.5	89.1±39	65.6±44.5	50.2±28.5	71.7±56.5
FEV1/FVC	64.7±4	58.2±18.5	44.9±20	37.1±17	48.5±23
DLCO (mL/min/mmHg)	49.3±16	48.6±25	44.3±15	31.2±11.5	41.7±27

BMI, body mass index; DLCO, diffusing capacity of the lung for carbon monoxide; FEV1%, forced expiratory volume in the first second (percentage of predicted); FEV1/FVC, forced expiratory volume in the first second over forced vital capacity; FVC%, forced vital capacity (percentage of predicted); GOLD,Global Initiative for Obstructing Lung Disease; Pk/y, packs per year.
Data are given as mean±standard deviation.

Intra- and interoperator variability study

Two residents (F.D., G.K.P.) with one year of experience in thoracic radiology were preliminarily trained for three months on the manual and semi-automated segmentation techniques and lobar volumetry, using chest CT scans unrelated to this study. After the three months dedicated to training, the same two operators started to analyze the CT scans of this study (one CT scan per patient). In a first run, both operators analyzed independently all of the 47 CT examinations, using both manual and semi-automated segmentations, as described below. At the end of the first run, the two operators suspended their analyses for three months. This interval of inactivity was selected according to previous clinical studies of lung volume reduction (17). After three months, the same two operators re-evaluated the same 47 CT examinations in a second run, using both segmentation procedures. In the two runs, both segmentation methods produced three-dimensional models of the lungs and lobes. The volumetry of these models was performed. The intra-and interoperator variability of pulmonary volumes and emphysema scores were calculated. The variability of volumes and emphysema scores obtained by semi-automated segmentation were compared with those obtained by manual segmentation.

Lung segmentation and volumetry

Manual segmentation

All of the CT examinations were analyzed on a clinical workstation using commercial software (Voxtool 6.7, GE Healthcare). The images were displayed on the orthogonal planes. Bone and soft tissues were excluded with a density mask of -1024/-200 Hounsfield unit (HU). The trachea, main bronchi, esophagus, and bowel were excluded by manual editing. The lungs were

also separated manually. To separate the lobes, each lung was displayed on the sagittal plane (Fig. 1). From the lateral to the medial aspect of the lung, all voxels outside the lower lobe were selected in a manually drawn region of interest and deleted. The remaining voxels of the lower lobe were saved as a three-dimensional model and subtracted from the native model of the lung. On the left side, this produced the model of the upper lobe. The right upper and middle lobes were separated by performing an additional segmentation with the same method. Manual segmentation required approximately 60 min per case.

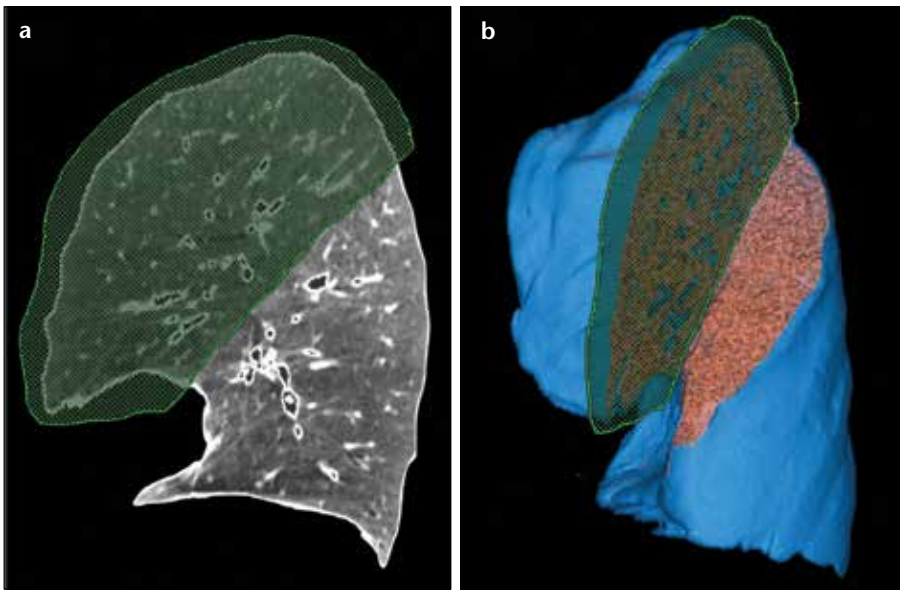

Figure 1. a, b. Lobar segmentation with the manual technique. Sagittal **(a)** and three-dimensional **(b)** views of the left lung during the segmentation of the lower lobe. Scrolling from the lateral to the medial aspect of the lung, the voxels representing the upper lobe were included in a region of interest (area in green) and deleted. At the end of the segmentation, the model of the lower lobe was completed and saved. The upper lobe was determined by subtraction of the lower lobe from the lung. The right upper and middle lobes were separated with an additional segmentation performed with the same method.

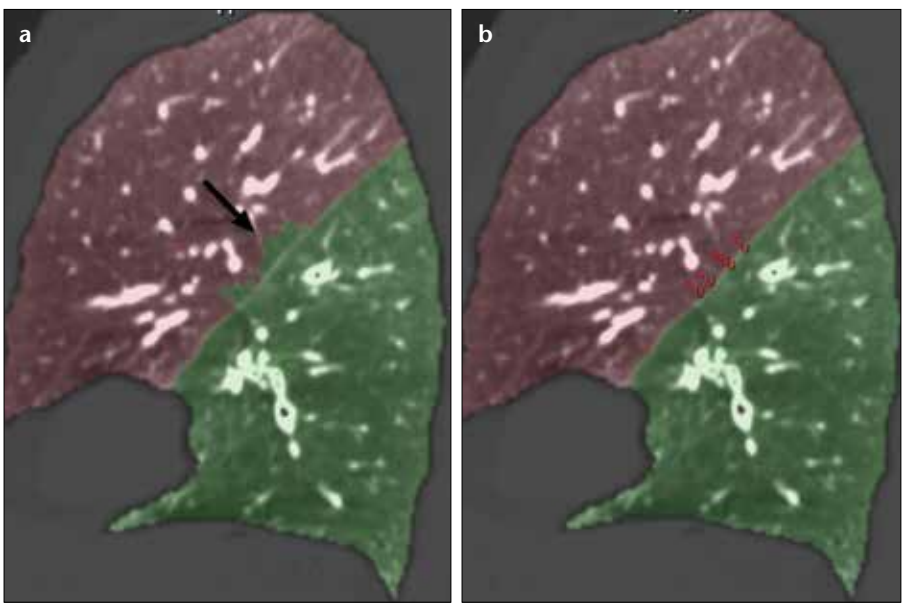

Figure 2. a, b. Lobar segmentation with the semi-automated technique. Automated segmentation delimited the outer borders of the lung and found the interlobar fissure (the upper lobe is color-coded in red, the lower lobe is color-coded in green), but incorrectly assigned a group of pixels of the lower lobe to the upper lobe (**a,** *arrow*). This area was marked with seed points identifying the upper lobe (crosses in red) and then automatically corrected by the software **(b)**.

Semi-automated segmentation

All of the CT examinations were also analyzed on a personal computer using research software (MeVisPULMO 3D® 3.071, Fraunhofer MEVIS, Bremen, Germany). The software excluded automatically bone and soft tissues using a density mask of -1024/-200 HU, separated the lungs from the tracheo-bronchial tree and other non pulmonary structures, and segmented the lobes (12). The lobes were identified using lung vessels as anatomic landmarks. Up to this point, the procedure was fully automated. Then, the correspondence between the lobes represented by CT and the superimposed color-coded areas that schematically identified them was reviewed. The operator reassigned the miscategorized areas of the lung using lobe-specific seeds. A representative case of this manual adjustment is shown in Fig. 2. The segmentation was reviewed until the lobes were correctly identified. Semi-automated segmentation required approximately 20 min per case, including the time for loading the CT images, the automated processing and revision of proposed segmentation with the manual refinements.

Lung volumetry

Both software programs automatically computed lung volumes by summing the volumes of voxels included in each anatomic model. The emphysema scores were determined as the percentage of segmented voxels having attenuation lower than eight thresholds ranging from -970 to -900 HU, in 10 HU increments, including thresholds validated by previous studies (18–20).

Statistical analysis

The intra- and interoperator variability of data obtained by manual and semi-automated segmentation was assessed with Bland-Altman analysis (21, 22). The coefficient of repeatability and the coefficient of reproducibility, representing 1.96 times the standard deviation of the differences between two sets of measurements, were calculated from the intraoperator (1 and 2) and inter-operator analyses, respectively (23). The variability of the volumes and emphysema scores obtained by semi-automated segmentation was compared with the variability

of the volumes and scores obtained by manual segmentation using the Maloney-Rastogi method (24). Holm-Bonferroni correction was applied for multiple comparisons. Corrected *P* values < 0.05 indicated statistically significant differences.

Results

The 47 CT examinations were suitable for quantitative analysis, and no variants of lobar anatomy or incomplete fissures were detected. The mean pulmonary volumes calculated using the reference method (i.e., manual segmentation) were 6066.7±1274.1 mL for the global lung volume, 3243.3±731.2 mL for the right lung, 2822.8±659.2 mL for the left lung, 1409.9±435.2 mL for the right upper lobe, 462.9±150 mL for the right middle lobe, 1369.7±369.8 mL for the right lower lobe, 1493.1±413.6 mL for the left upper lobe, and 1329.3±387.9 mL for the left lower lobe.

Variability of global and single lung volumes

In the intraoperator analyses, the coefficients of repeatability of volumes obtained by manual segmentation of the global, right, and left lungs were 65 mL, 40.9 mL, 25.1 mL for the first operator and 16.7 mL, 9.5 mL, 9.6 mL for the second operator, respectively. In the interoperator analysis, the coefficients of reproducibility of volumes obtained by manual segmentation of global, right, and left lungs were 81.9 mL, 49.9 mL, and 38.3 mL, respectively. Lung segmentation performed with the research software in fully automated mode yielded no variability for global, right, and left lung volumes.

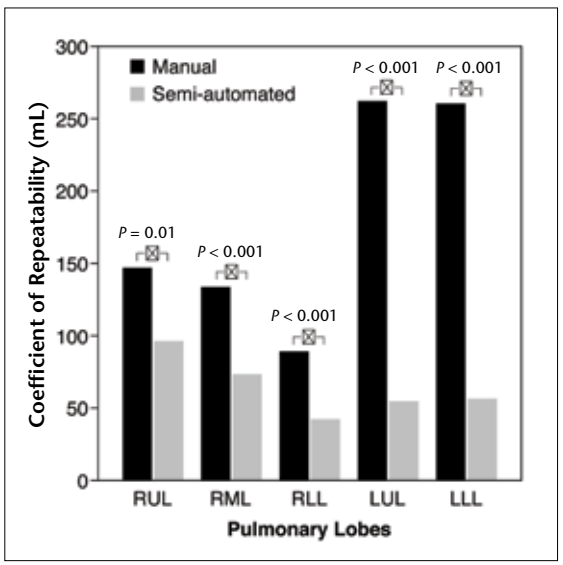

Figure 3. Variability of lobar pulmonary volumes calculated from the intraoperator analysis of the first operator. The coefficients of repeatability of lobar volumes calculated by semi-automated segmentation were lower than those from the manual segmentation. RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL, left upper lobe; LLL, left lower lobe.

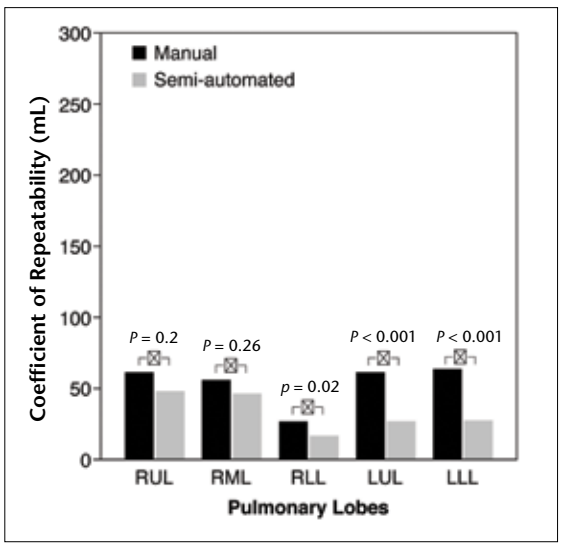

Figure 4. Variability of lobar pulmonary volumes calculated from the intraoperator analysis of the second operator. The coefficients of repeatability of lobar volumes calculated by semi-automated segmentation were lower than from the manual segmentation. RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL, left upper lobe; LLL, left lower lobe.

Variability of lobar volumes

In the intraoperator analyses, the coefficients of repeatability of lobar volumes calculated by semi-automated lobe segmentation were lower than those from manual segmentation (first operator in Fig. 3, second operator in Fig. 4). In the interoperator analysis, the coefficients of reproducibility of lobar volumes calculated by semi-automated lobe segmentation were also lower than those from manual segmentation (Fig. 5).

Variability of emphysema scores of global and single lungs

In the intraoperator analyses, the coefficients of repeatability of the emphysema scores obtained at different thresholds by manual segmentation of global and single lungs were in the range of 0.06%–2.29% for the first operator and 0.07%–1.75% for the second operator, respectively. In the interoperator analysis, the coefficient of reproducibility of emphysema scores obtained at different thresholds by manual segmentation of global and single lungs was in the range of 0.13%–2.27%. As for the volume, lung segmentation performed with the research software in fully automated mode yielded no variability of emphysema scores for global, right, and left lungs.

Variability of emphysema scores of the lobes

Table 2 summarizes the ranges of coefficients of repeatability and of reproducibility of the emphysema scores computed at different thresholds for the pulmonary lobes. In the intraoperator analysis of the first operator, the difference between the coefficients of repeatability obtained by manual and semi-automated lobe segmentation for the middle lobe was significant at the thresholds of -930 and -970 HU ($P < 0.001$ and $P = 0.002$, respectively) and not significant for other thresholds. In the intraoperator analysis of the second operator, the difference between the coefficients of repeatability obtained by manual and semi-automated segmentation was not significant for the right upper lobe at the threshold of -940 HU and -960 HU ($P = 0.31$ and $P = 0.073$, respectively). In the interoperator analysis, the difference between the coef-

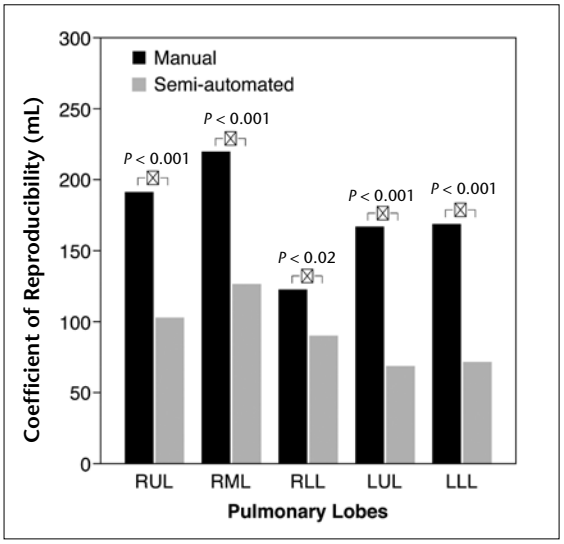

Figure 5. Variability of lobar pulmonary volumes calculated from the interoperator analysis. The coefficients of reproducibility of lobar volumes calculated by semi-automated segmentation were lower than those by manual segmentation. RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL, left upper lobe; LLL, left lower lobe.

Table 2. Variability of the emphysema scores: manual vs. semi-automated segmentation

	Manual	Semi-automated	Corrected *P*
Intraoperator 1			
Right upper lobe	1.17–1.72	0.50–0.69	0.002
Right middle lobe	3.03–10.1	1.83–3.21	< 0.001–0.59[a]
Right lower lobe	1.10–2.21	0.59–1.02	0.003
Left upper lobe	1.14–4.13	0.27–0.61	< 0.001
Left lower lobe	0.41–1.93	0.31–0.34	0.049
Intraoperator 2			
Right upper lobe	0.41–3.33	0.27–0.58	0.027[b]
Right middle lobe	1.79–4.56	0.73–1.58	0.039
Right lower lobe	0.65–8.57	0.20–0.41	0.001
Left upper lobe	0.25–2.45	0.09–0.24	0.011
Left lower lobe	0.16–0.35	0.08–0.19	0.031
Interoperator			
Right upper lobe	1.36–2.04	0.64–1.13	0.055
Right middle lobe	4.75–6.29	2.42–4.30	0.023[c]
Right lower lobe	0.60–2.08	0.23–0.74	0.036
Left upper lobe	1.02–2.85	0.25–0.78	< 0.001
Left lower lobe	0.43–1.70	0.31–0.61	0.003[c]

Data in table represent ranges of coefficients of repeatability (intraoperator analyses) and of reproducibility (interoperator analysis) of emphysema scores (%) obtained at eight density thresholds (-900, -910, -920, -930, -940, -950, -960, -970 HU). Reported in the last column are the highest *P* values obtained from individual comparisons between manual and semi-automated methods at each threshold, unless differently indicated (Maloney-Rastogy test).

[a]*P* values were significant at the density thresholds of -930 and -970 HU (< 0.001 and 0.002, respectively).

[b]*P* values were not significant at -940 HU and at -960 HU (0.31 and 0.073, respectively).

[c]*P* values were not significant at -900 HU for the right middle lobe and at -970 HU for the left lower lobe (0.091 and 0.071, respectively).

ficients of reproducibility obtained by manual and semi-automated segmentation was not significant for the right middle lobe at the threshold of -900 HU and for the left lower lobe at the threshold of –970 HU ($P = 0.091$ and $P = 0.071$, respectively). In all other analyses, the coefficients of repeatability and reproducibility of the emphysema scores decreased significantly when lobar volumetry was performed by semi-automated segmentation.

Discussion

In this study, we demonstrated that semi-automated segmentation reduces the intra- and interoperator variability of lobar volumetry in lung emphysema. Indeed, the coefficients of repeatability and of reproducibility of lobar volumes and emphysema scores calculated by the semi-automated segmentation were lower than those of the manual method. Using the semi-automated segmentation technique, the interoperator variability of the lobar volumes and emphysema scores ranged from ±68.7 to ±126.5 mL and from ±0.61% to 4.30%, respectively, depending on the target lobe and selected threshold of emphysema.

Surgical or bronchoscopic lung volume reduction is a therapeutic option for patients with advanced emphysema (3). When emphysema affects more than 50% of the volume of one lobe and relatively spares adjacent ipsilateral lobes, the patient can benefit from bronchoscopic lung volume reduction (3). CT quantifies the severity and distribution of the disease and monitors the effects of treatment (3, 4). However, the intrinsic variability of CT measures of lobar volumes and emphysema scores can potentially influence the therapeutic management of these patients (5). We analyzed the intra- and interoperator variability of lobar volumes and emphysema scores in a group of COPD patients with emphysema that had been investigated with CT in a clinical setting. The retrospective method of the study was supported by previous investigation (25). All of the chest CT examinations were technically identical and presented no signs of anatomic variants that could prevent or influence lung segmentation and volumetry.

A preliminary step for quantitative CT analysis of lung volume and attenuation is to select an anatomic target by image segmentation (12, 26). Manual segmentation of voxels representing an organ or tissue of interest is a time-consuming procedure. However, the manual technique is still the standard reference for validating new emerging and more advanced computer-based segmentation algorithms (12, 15, 26). We compared the variability of this manual approach to that of a software-based algorithm, which automatically separates the pulmonary

lobes using lung vessels as anatomic landmarks and requires interactive modifications for segmentation errors. The manual method was performed on a clinically approved software platform. The semi-automated software method was previously validated (27). Both methods used the same density thresholds (-1024/-200 HU) for preprocessing of image data that aimed to separate the lungs from bone and soft-tissues.

As lung vessels are altered in advanced emphysema (28), we reviewed all segmentations performed by the software-based automated approach. The operators refined the segmented lobes by creating seed points on the vascular markings. The lung territory depending on those vessels was automatically re-assigned to the correct lobe. This human interaction was sufficient to consider variability in the measurements of lobar volumes and emphysema scores. Both parameters were calculated with the same procedure using both manual and semi-automated methods. The emphysema scores, which indicated the severity of the disease in each anatomic target (3), were calculated at multiple thresholds in a range that included thresholds validated by previous studies (18). The range of variability of the scores imposed by using multiple thresholds was assessed. Manual removal of trachea, main bronchi, and other nonpulmonary structures determined the variability of the global and single lung volumes. Automated segmentation produced no variability for these anatomic targets. For the lobar volumes, variability due to manual segmentation has been reported previously (10). In our study, semi-automated segmentation reduced this variability in both intra- and interoperator analyses. The semi-automated method was particularly beneficial in reducing the variability of lobar volumes calculated by the first operator, whose results with the manual approach appeared more variable than those of the second operator.

In all analyses, the left upper and lower lobes had almost identical variability. Those lobes share the same source of variability, which is the delineation of one fissure as boundary for two lobes. On the right side, the lower lobe presented lower volume variability compared to the upper and middle lobes. Again, the variability for the right lower lobe depends on the segmentation along one fissure (the major one). The right upper and middle lobes share an additional source of variability, which is the segmentation on the other fissure (the minor one). This may partly explain the nonsignificant difference between the variability of the volumes of the right upper and middle lobes between the manual and semi-automated methods, respectively, for the second operator (Fig. 4). Those differences might be nonsignificant either because the second operator produced low variability of lobar volumes manually or because the semi-automated approach with segmentation involving two fissures did not sufficiently reduce the variability of those volumes. A trend of high variability was also noted for the right middle lobe in the interoperator analysis. As reported previously, the smaller absolute volume of the right middle lobe leads to considerable relative differences, which are also expected because the minor fissure is typically less pronounced and therefore more difficult to delineate (10).

As for the lung volumes, the automated method completely removed variability of emphysema scores calculated at different thresholds for the global and single lungs. In all lobar analyses, the semi-automated method reduced the variability of the emphysema scores. At some density thresholds, the differences of variability of emphysema scores were not significant, and no relationship could be established with the corresponding differences of variability of lobar volumes. However, the variability of emphysema scores is not completely dependent on the variability of total volume. The repeated segmentation of a pulmonary lobe may produce identical volumes but different density histograms, depending on the voxels included in the segmentation. As a fraction of the total volume calculated in a fixed density range, the emphysema score has variability that depends on the variability of the total volume and the variability of the histogram shape.

Other studies have recently investigated the variability of lobar volumetry. Brown et al. (8) evaluated the variability of volume estimates between two CT examinations performed at baseline and after an interval of nine months (long-term interscan variability) in a large study group including patients who are candidates for bronchoscopic volume reduction. Chong et al. (9) performed a similar study evaluating the short-term interscan variability after an interval of one week between two CT scans. The interscan variability reflects differences between subsequent scans in the same subject caused by scanner instability and physiological differences (e.g., inspiration level) (5). Our study focused on the intra- and interoperator variability of pulmonary volumes, which reflects differences between analyses on the same data performed by the same (intra-) or different (inter-) operator (5). Recently, Revel et al. (27) investigated the automated quantification of emphysema in patients with COPD. They compared the lobar emphysema scores obtained by refined segmentation with those by unrefined segmentation and evaluated the feasibility of the fully automated approach. Biases were negligible and the intervals of agreement of lobar volumes between the two approaches were in the range of ±7% to ±30%. We compared the intra- and interoperator variability of lobar volumes and emphysema scores obtained by manual segmentation with those by semi-automated segmentation with operator-dependent editing (comparison between variabilities of the two methods).

This study has several limitations. We assessed retrospectively CT examinations performed for clinical evaluation of emphysema, which does not routinely include dedicated calibration of the scanner for air and blood density (5) and monitoring of patient inspiration level. In addition, volume correction of lung densitometry was also not performed (29). These factors did not influence our analysis of intra- and interoperator variability that included one CT examination per patient. However, we could not determine the inter subject variability in our study group. The study included only patients with COPD and no control groups, which means that potential differences of

variability induced by emphysema were not determined. The measurements of variability were also not performed according to emphysema severity. Variants of fissure and lobar anatomy were not present in this patient series; therefore, no conclusion can be drawn in terms of differences of variability in presence of incomplete fissures and accessory lobes. A study in COPD patients with severe emphysema has recently shown that the interobserver agreement in visually detecting the presence of incomplete fissures can be reduced if the evaluation is performed by experienced radiologists (30). Additionally, a method for investigating the completeness of pulmonary fissures has been proposed (13). The semi-automated segmentation was evaluated using one software program. A difference in terms of volume between the automated segmentation and the semi-automated segmentation obtained after refinements was not available. Therefore, we are not able to provide an estimate of the error size of the fully automated technique. The variability of other CT measures of emphysema such as the 15th percentile point of lung density could not be assessed for differences between the two software programs. Comparison studies should also assess potential differences among different semi-automated segmentation techniques. Finally, an *in vivo* gold standard for determination of lobar volumes is not currently available. This factor did not influence the measurements of the variability of the lobar volumetry. However, to validate the accuracy of the technique, phantom studies should be performed.

In conclusion, semi-automated segmentation reduces the intra- and interoperator variability of lobar volumetry and provides a more objective tool to quantify lung volumes and severity of emphysema.

Conflict of interest disclosure

The authors declared no conflicts of interest.

References

1. World Health Organization Report on the Global Tobacco Epidemic: Warning about the dangers of tobacco. World Health Organization, Geneva 2011. http://whqlibdoc.who.int/publications/2011/9789240687813_eng.pdf. Accessed on September 27, 2011.
2. Strange C, Herth FJF, Kovitz KL, et al. Design of the Endobronchial Valve for Emphysema Palliation Trial (VENT): a non-surgical method of lung volume reduction. BMC Pulm Med 2007; 7:10. [CrossRef]
3. Sciurba FC, Ernst A, Herth FJF, et al. A randomized study of endobronchial valves for advanced emphysema. N Engl J Med 2010; 363:1233–1244. [CrossRef]
4. Brown MS, Kim HJ, Abtin FG, et al. Emphysema lung lobe volume reduction: effects on the ipsilateral and contralateral lobes. Eur Radiol 2012; 22:1547–1555. [CrossRef]
5. Bakker ME, Stolk J, Putter H, et al. Variability in densitometric assessment of pulmonary emphysema with computed tomography. Invest Radiol 2005; 40:777–783. [CrossRef]
6. Sverzellati N, Molinari F, Pirronti T, Bonomo L, Spagnolo P, Zompatori M. New insights on COPD imaging via CT and MRI. Int J Chron Obstruct Pulmon Dis 2007; 2:301–312.
7. Bankier AA, De Maertelaer V, Keyzer C, Gevenois PA. Pulmonary emphysema: subjective visual grading versus objective quantification with macroscopic morphometry and thin-section CT densitometry. Radiology 1999; 211:851–858.
8. Brown MS, Kim HJ, Abtin F, et al. Reproducibility of lung and lobar volume measurements using computed tomography. Acad Radiol 2010; 17:316–322. [CrossRef]
9. Chong D, Brown MS, Kim HJ, et al. Reproducibility of volume and densitometric measures of emphysema on repeat computed tomography with an interval of 1 week. Eur Radiol 2012; 22:287–294. [CrossRef]
10. Molinari F, Amato M, Stefanetti M, et al. Density-based MDCT quantification of lobar lung volumes: a study of inter- and intraobserver reproducibility. Radiol Med 2010; 115:516–525. [CrossRef]
11. Pham DL, Xu C, Prince JL. Current methods in medical image segmentation. Annu Rev Biomed Eng 2000; 2:315–337. [CrossRef]
12. Kuhnigk J-M, Dicken V, Zidowitz S, et al. Informatics in radiology (infoRAD): new tools for computer assistance in thoracic CT. Part 1. Functional analysis of lungs, lung lobes, and bronchopulmonary segments. Radiographics 2005; 25:525–536. [CrossRef]
13. van Rikxoort EM, Goldin JG, Galperin-Aizenberg M, et al. A method for the automatic quantification of the completeness of pulmonary fissures: evaluation in a database of subjects with severe emphysema. Eur Radiol 2012; 22:302–309. [CrossRef]
14. Sluimer I, Prokop M, van Ginneken B. Toward automated segmentation of the pathological lung in CT. IEEE Trans Med Imaging 2005; 24:1025–1038. [CrossRef]
15. van Rikxoort EM, Prokop M, de Hoop B, Viergever MA, Pluim JP, van Ginneken B. Automatic segmentation of the pulmonary lobes from fissures, airways, and lung borders: evaluation of robustness against missing data. Med Image Comput Comput Assist Interv 2009; 12:263–271.
16. Sverzellati N, Calabro E, Randi G, et al. Sex differences in emphysema phenotype in smokers without airflow obstruction. Eur Respir J 2009; 33:1320–1328. [CrossRef]
17. Emery MJ, Eveland RL, Eveland K, Couetil LL, Hildebrandt J, Swenson ER. Lung volume reduction by bronchoscopic administration of steam. Am J Respir Crit Care Med 2010; 182:1282–1291. [CrossRef]
18. 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.
19. Gierada DS, Bierhals AJ, Choong CK, et al. Effects of CT section thickness and reconstruction kernel on emphysema quantification relationship to the magnitude of the CT emphysema index. Acad Radiol 2010; 17:146–156. [CrossRef]
20. Madani A, Zanen J, de Maertelaer V, Gevenois PA. Pulmonary emphysema: objective quantification at multi-detector row CT--comparison with macroscopic and microscopic morphometry. Radiology 2006; 238:1036–1043. [CrossRef]
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307–310. [CrossRef]
22. Halligan S. Reproducibility, repeatability, correlation and measurement error. Br J Radiol 2002; 75:193–195.
23. Barnhart HX, Haber MJ, Lin LI. An overview on assessing agreement with continuous measurements. J Biopharm Stat 2007; 17:529–569. [CrossRef]
24. Krummenauer F, Genevrière I, Nixdorff U. The biometrical comparison of cardiac imaging methods. Comput Methods Programs Biomed 2000; 62:21–34. [CrossRef]
25. Hersh CP, Washko GR, Jacobson FL, et al. Interobserver variability in the determination of upper lobe-predominant emphysema. Chest 2007; 131:424–431. [CrossRef]
26. Leader JK, Zheng B, Rogers RM, et al. Automated lung segmentation in X-ray computed tomography: development and evaluation of a heuristic threshold-based scheme. Acad Radiol 2003; 10:1224–1236. [CrossRef]
27. Revel MP, Faivre JB, Remy-Jardin M, et al. Automated lobar quantification of emphysema in patients with severe COPD. Eur Radiol 2008; 18:2723–2730. [CrossRef]
28. Matsuoka S, Washko GR, Dransfield MT, et al. Quantitative CT measurement of cross-sectional area of small pulmonary vessel in COPD: correlations with emphysema and airflow limitation. Acad Radiol 2010; 17:93–99. [CrossRef]
29. Shaker SB, Dirksen A, Laursen LC, Skovgaard LT, Holstein-Rathlou NH. Volume adjustment of lung density by computed tomography scans in patients with emphysema. Acta Radiol 2004; 45:417–423. [CrossRef]
30. Koenigkam-Santos M, Puderbach M, Gompelmann D, et al. Incomplete fissures in severe emphysematous patients evaluated with MDCT: Incidence and interobserver agreement among radiologists and pneumologists. Eur J Radiol 2012; 81:4161–4166. [CrossRef]