IOP PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **53** (2008) 1295–1312 doi:10.1088/0031-9155/53/5/009

Effect of CT scanning parameters on volumetric measurements of pulmonary nodules by 3D active contour segmentation: a phantom study

Ted W Way, Heang-Ping Chan, Mitchell M Goodsitt, Berkman Sahiner, Lubomir M Hadjiiski, Chuan Zhou and Aamer Chughtai

Department of Radiology, University of Michigan, Ann Arbor, MI 48109, USA

E-mail: tway@umich.edu

Received 17 September 2007, in final form 20 December 2007
Published 12 February 2008
Online at stacks.iop.org/PMB/53/1295

Abstract

The purpose of this study is to investigate the effects of CT scanning and reconstruction parameters on automated segmentation and volumetric measurements of nodules in CT images. Phantom nodules of known sizes were used so that segmentation accuracy could be quantified in comparison to ground-truth volumes. Spherical nodules having 4.8, 9.5 and 16 mm diameters and 50 and 100 mg cc^{-1} calcium contents were embedded in lung-tissue-simulating foam which was inserted in the thoracic cavity of a chest section phantom. CT scans of the phantom were acquired with a 16-slice scanner at various tube currents, pitches, fields-of-view and slice thicknesses. Scans were also taken using identical techniques either within the same day or five months apart for study of reproducibility. The phantom nodules were segmented with a three-dimensional active contour (3DAC) model that we previously developed for use on patient nodules. The percentage volume errors relative to the ground-truth volumes were estimated under the various imaging conditions. There was no statistically significant difference in volume error for repeated CT scans or scans taken with techniques where only pitch, field of view, or tube current (mA) were changed. However, the slice thickness significantly ($p < 0.05$) affected the volume error. Therefore, to evaluate nodule growth, consistent imaging conditions and high resolution should be used for acquisition of the serial CT scans, especially for smaller nodules. Understanding the effects of scanning and reconstruction parameters on volume measurements by 3DAC allows better interpretation of data and assessment of growth. Tracking nodule growth with computerized segmentation methods would reduce inter- and intraobserver variabilities.

1. Introduction

Computed tomography (CT) is more sensitive than chest radiography in detecting small and subtle lung nodules (Henschke *et al* 1999, Diederich *et al* 2002, Kaneko *et al* 1996, Nawa *et al* 2002, Sone *et al* 1998, 2001, Swensen *et al* 2005, 2003). With the advent of multi-detector row helical CT scanners that offer ever-thinner slices, even more nodules are detected (Fischbach *et al* 2003). As it is recommended to follow up on most (MacMahon *et al* 2005), if not all (Ost *et al* 2003, Tan *et al* 2003) detected nodules, the management of these nodules may overwhelm radiologists, especially if lung cancer screening with CT becomes recommended practice.

Computer-aided diagnosis (CAD) tools have the potential to assist radiologists as a second reader in detecting, classifying (characterizing) and managing lung nodules. Currently, a CAD system can analyse a single scan, segmenting and then extracting features from the nodule to estimate its likelihood of malignancy. Features such as morphology (volume, size and shape) (Way *et al* 2006, Armato *et al* 2003, Kawata *et al* 1998), histograms of grey levels, the contrast of the nodule (Li *et al* 2004), enhancement after contrast injection (Shah *et al* 2005) and texture (Way *et al* 2006) have been used. There has also been increasing attention given to volume doubling rate as an indicator of malignancy (Lillington and Caskey 1993, Usuda *et al* 1994).

To track volume growth, an accurate boundary of the nodule is needed, and it is challenging for many reasons. First, it is difficult to determine nodule boundaries definitively, as suggested by the considerable inter- and intra-observer variabilities among radiologists in assessing lung nodule size (Bogot *et al* 2005, Erasmus *et al* 2003, Meyer *et al* 2006). Second, the method used to measure volume may vary. Nodule size can only be measured in two dimensions on a CT slice, typically based on its longest and perpendicular axes. It is difficult to assess doubling time, especially for smaller nodules, for which a small change in diameter corresponds to a large change in the volume. For example, a sphere with a 3 mm diameter doubles in volume when the diameter increases by only 26%, to 3.78 mm. This 0.78 mm increase is about the size of 1.1 pixels in a typical scan with a field of view (FOV) of 36 cm.

Different window width and level settings can also affect the measurement of diameters (Harris *et al* 1993). These factors contribute to inaccuracies in the measured volume (Winer-Muram *et al* 2003), resulting in inconsistency and uncertainty in detecting volume change in serial CT scans. The large inter- and intra-observer variations in radiologists' manual segmentation demonstrated in the 23 nodules provided by the lung image database consortium (LIDC) highlight the difficulty in judgment of nodule boundaries on CT scans (Meyer *et al* 2006). Yankelevitz *et al* found that computerized methods are accurate on phantom nodules and applied the methods to clinical patient scans over time. They concluded that computerized methods using CT volumetric information that are not affected by windowing settings or intra- and inter- observer errors will be useful for nodule volume estimation and for assessing growth (Yankelevitz *et al* 1999, 2000).

Even though computerized methods may be immune to some factors that may influence radiologists, such as variation in windowing, there are still challenges due to image acquisition parameters (Zerhouni *et al* 1982, Im *et al* 1988). Various groups have investigated the effects of image acquisition parameters using nodule-mimicking spheres with known volumes. Ko *et al* reported computer-calculated volumes from regions of interest (ROIs) marked by radiologists. Using a threshold method for segmentation, they found that tube current-time (20 and 120 mAs), reconstruction algorithm, quantitative volume calculation method, nodule attenuation (solid at 50 HU and ground glass at −360 HU) and nodule size significantly affected the volume error (Ko *et al* 2003). Goo *et al* measured volumes using a thresholding

method and reported that section thickness and threshold significantly affected absolute volume error (Goo *et al* 2005).

The major differences between this investigation and previous studies (Harris *et al* 1993, Ko *et al* 2003, Winer-Muram *et al* 2003, Goo *et al* 2005) include the segmentation algorithm, the greater ranges of scanning and reconstruction parameters and nodule sizes, and the effect of calcium concentration of the nodules. In addition, we investigated the reproducibility of repeated scans performed either consecutively or separated by a few months (five in this study) as in clinical follow-up studies. An automated nodule segmentation method was used so the volume errors were assessed free of inter- and intra-observer variabilities.

Although phantom nodules are different from real pulmonary nodules in many ways, phantom studies will allow comparison with the ground-truth volumes and systematic analysis of the relative trends of the volume error dependence on CT imaging conditions. This would be virtually impossible to perform with nodules in patient scans. Besides dose concerns, motion artefacts can be a major variable, making it difficult to determine what actually contributes to the measured volume errors. The fixed size of chest phantoms also precludes nodule volume variations due to lung volume change when breathing (Petkovska *et al* 2007). Despite the fact that the absolute volume errors estimated with phantoms will not be applicable to real nodules, the results will provide useful information for the selection of CT imaging protocols for clinical follow-up and for the interpretation of nodule volume changes.

2. Methods and materials

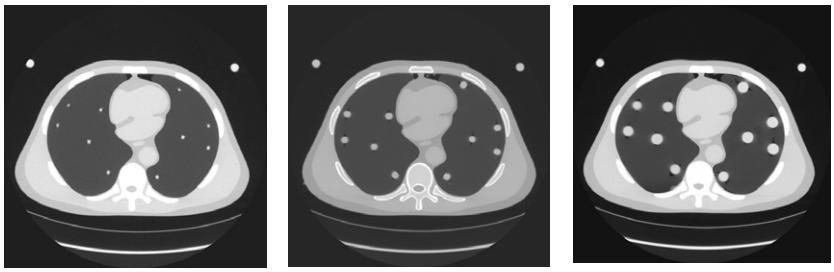
2.1. Chest phantom and spherical nodules

The chest phantom consisted of a CIRS Model 003 tissue equivalent transaxial thorax section phantom (CIRS, Norfolk, VA) sandwiched between two large water-equivalent bolus sections (Goodsitt *et al* 2006). The body shape and lung cavities of the water bolus sections match those of the thorax section. The lung cavities in the thorax section were filled with a foam that has similar CT number Hounsfield Units (HU) as lung tissue in CT scans (Goodsitt *et al* 2006). Nodule-mimicking spheres of diameters 4.8 mm, 9.5 mm and 16 mm were scanned to evaluate the dependence on the nodule size. These spheres were made of an epoxy resin with added fillers that would produce x-ray linear coefficients nearly identical to that as water in the CT energy range. The amount of calcium carbonate ($CaCO_3$) added to the spheres determined their calcium density. These sizes were chosen by taking into consideration the nodule sizes that are of clinical significance (3 to 30 mm) and the expected larger errors in volume estimation for smaller nodules in CT scans. They were inserted in slits cut in the foam of the lung section. The interface between sections was filled with Vaseline to minimize air gaps.

In experiment I, five 4.8 mm diameter 50 mg cc^{-1} $CaCO_3$ spheres were placed in the left lung cavity, and six 4.8 mm diameter 100 mg cc^{-1} $CaCO_3$ spheres were placed in the right lung cavity. In experiment II, scans taken five months later, 10 nodules of the same size were placed in the thorax section phantom. Five $CaCO_3$ spheres of 50 mg cc^{-1} were placed in the right lung cavity, and five 100 mg cc^{-1} spheres were placed in the left lung cavity. After scanning was completed for one size (e.g. 4.8 mm), all ten spheres were replaced with those of another size (e.g. 9.5 mm). Examples of the CT scans acquired are shown in figure 1.

2.2. Reproducibility of lung nodule volume measurement

CT scans were acquired with a GE LightSpeed VCT 64 slice scanner operating in 20 mm collimation mode (GE Healthcare, Waukesha, WI). Images were acquired in two sessions



(1) **(2)** **(3)**

Figure 1. Examples of two-dimensional (2D) axial slices for the three phantom nodule configurations used in experiment II: the nodules have diameters of (1) 4.8 mm, (2) 9.6 mm and (3) 16 mm.

(experiments I and II) five months apart. Within each experiment, three scans (scans 1, 2 and 3) using identical parameters were taken for each imaging condition. The same imaging conditions were used in experiments I and II to simulate serial exams where a follow-up scan would be taken months later. Thus, variability between the two experiments and within each experiment can be assessed.

We focused on protocols used with GE 16-slice scanners in two clinical trials. These conditions were employed because neither trial had published protocols for the new 64-slice scanners at the time we initiated our study, and the 64-slice scanner could be operated in a 20 mm collimation instead of 40 mm collimation mode, which is similar to that of the 16-slice scanners. The first source of parameters was the national lung screening trial (NLST), sponsored by the National Cancer Institute (NCI), which enrolled over 50 000 high-risk subjects to compare the effectiveness of CT and chest radiography in lung cancer detection. Scans on GE 16-slice scanners were acquired with the 16 $\times$ 1.25 mm detector configuration, 120 kVp, 80 mA, 0.5 s rotation time and 1.375:1 pitch. The images were reconstructed to 1.25 mm slice thickness and 1.25 mm slice interval for a thin slice data set, in addition to the 2.5 mm slice thickness and 2.0 mm slice interval data set. The second source was the lung tissue resource consortium (LTRC), sponsored by the National Heart, Lung and Blood Institute, the goal of which was to create a lung tissue database to understand the pathogenetic mechanisms of lung diseases. CT scans were collected as part of the process. Scans on GE 16-slice scanners were acquired at 140 kVp, 300 mA and 0.5 s rotation time, 1.375:1 pitch. The data images were reconstructed at 0.625 mm slice thickness and 0.625 mm slice interval, and a retrospective reconstruction was also done at 1.25 mm thickness and 0.625 mm interval. Since both are large-scale studies, and participants in the NLST underwent multiple scans over time, accuracy and reproducibility under these CT protocols are of interest. An additional high-resolution imaging condition was chosen in anticipation of thin-slice scans that may become more common with multi-detector row CT scanners. All scans were reconstructed with the standard kernel in the GE scanners. The imaging conditions for the reproducibility study are listed in table 1.

2.3. Effects of CT parameters on volume errors

For the study of the effects of CT scanning and reconstruction parameters on nodule volume estimation, each parameter was varied in a range that may be used in clinical examinations, including the pitch (0.531:1 to 1.375:1), slice thickness (0.625 mm to 2.5 mm), tube current (80 mA to 400 mA) and field of view (25 cm to 36 cm).

Table 1. Table of scanning and reconstruction parameters for reproducibility study, which included the techniques used in the NLST and LTRC protocols and a high-resolution technique. In both protocols, an initial reconstruction and a retrospective reconstruction are performed from the original projection data.

Protocol	Pitch	Slice thickness (mm)	Slice interval (mm)	kVp	mA	Time (s)
High-resolution	0.531:1	0.625	0.625	120	400	0.8
NLST initial	1.375:1	1.25	1.25	120	80	0.5
NLST retrospective	1.375:1	2.50	2.00	120	80	0.5
LTRC initial	1.375:1	0.625	0.625	140	300	0.5
LTRC retrospective	1.375:1	1.250	0.625	140	300	0.5

2.4. Nodule segmentation

2.4.1. Image pre-processing. A CT scan input to the segmentation program is first pre-processed to obtain isotropic voxels. Linear interpolation in the z-axis is performed if the reconstructed slice interval of the CT scan is greater than the pixel size on the axial plane. Bilinear interpolation in the axial plane is used instead if the slice interval is smaller. The interpolation does not improve the spatial resolution of the CT data, but it is performed to facilitate the implementation of the three-dimensional (3D) segmentation operations.

Using a graphical user interface, each nodule location is manually identified by setting a box that contains the nodule on one scan based on visual inspection. For other scans where the nodules are in the same positions, an automatic nodule detection program extracts volumes of interest (VOIs) from the scan (Gurcan *et al* 2002, Ge *et al* 2005). The automatic detection is not a part of this study but is used to reduce the time of otherwise manually marking VOIs. For nodules that are close to the pleura, the VOI may include voxels from the pleura and the chest wall. For these nodules, a lung region-masking program previously developed in our laboratory (Gurcan *et al* 2002) is applied to the VOI to exclude these voxels from further processing.

2.4.2. Parametric active contour (AC) segmentation. Deformable contour models, particularly the AC model (Kass *et al* 1987), are well-known tools for image segmentation. ACs are energy-minimizing splines guided by various forces, or energies. We implemented a parametric model, representing the contour as vertices of a polygon, and minimized it based on a greedy algorithm described by Williams and Shah (1992). We implemented two new energies to take advantage of 3D volumetric data and one energy to penalize the total energy if the contour grows into the chest wall and mediastinum (Way *et al* 2006). The energy function is shown in equation (1):

$$E_{\text{total}} = \min_{E(c)} \sum_{c=1}^{N} [w_{\text{hom}} E_{\text{hom}}(c) + w_{\text{cont}} E_{\text{cont}}(c) + w_{\text{curv}} E_{\text{curv}}(c) + w_{3\text{D}_{\text{curv}}} E_{3\text{D}_{\text{curv}}}(c) + w_{\text{grad}} E_{\text{grad}}(c) + w_{3\text{D}_{\text{grad}}} E_{3\text{D}_{\text{grad}}}(c) + w_{\text{bal}} E_{\text{bal}}(c) + w_{\text{mask}} E_{\text{mask}}(c)], \quad (1)$$

where $E(c)$ is the energy at a pixel in the search neighbourhood of vertex $v(c)$. In this energy function, the internal energies include homogeneity (hom), continuity (cont), curvature (curv), 3D curvature (3D_{curv}), and the external energies include gradient (grad), 3D gradient (3D_{grad}), balloon (bal) and mask (mask). The weight w_j is a parameter assigned to each energy j, where j represents one of the eight energies: hom, cont, curv, 3D_{curv}, grad, 3D_{grad}, bal and mask. Details of how the weights were trained have been described elsewhere (Way *et al* 2006).

In this study, we applied the 3DAC model trained on patient nodules (Way *et al* 2006) to the segmentation of phantom nodules on CT scans. An initial contour for the nodule in the VOI is generated using a *k*-means clustering and object identification method and is used to initialize the 3DAC as explained in detail in Way *et al* (2006).

2.5. Manual segmentation by a radiologist

To provide a reference for comparison with the 3DAC contours, a fellowship-trained thoracic radiologist manually segmented some 4.8 mm diameter nodules. We selected three scanning conditions, which were identical except for slice thickness values of 0.625, 1.25 and 2.5 mm. The other parameters were 1.375:1 pitch, 0.625 mm slice interval, 120 kVp, 160 mA, 36 cm FOV and 0.5 s rotation time. The radiologist segmented the five 4.8 mm diameter nodules with 100 mg cc^{-1} density in the right lung of the chest phantom. There were a total of 15 manually segmented nodules from the five nodules and three different slice thickness conditions.

The radiologist was only informed that the phantom nodules were spherical, without being given any size information. A graphical-user interface (GUI) program developed in our laboratory displayed a zoomed in ROI containing the nodule to be segmented. The radiologist was free to adjust the brightness and contrast of the image. He was instructed to outline the nodule by a polygon following where he judged to be the nodule boundary in the image.

2.6. Data analysis

The volume of a segmented phantom nodule (V_{3DAC}) was calculated by multiplying the number of voxels contained in the contours with the volume of each voxel. The segmented volume was compared with the ground-truth volume of the phantom nodule (V_{true}) and the difference was reported as the percentage volume error given in equation (2):

$$\%\ \text{Vol err} = \frac{V_{3DAC} - V_{true}}{V_{true}} \times 100\%, \tag{2}$$

where V_{true} is calculated by the volume formula of a sphere, $\text{Vol} = (4/3)\pi r^3$. The volume error percentage is a signed value: positive for overestimation and negative for underestimation of the true volume.

The volume errors were calculated for all nodule sizes imaged under various CT imaging conditions. We estimated statistical significance based on Student's *t*-test for paired data and one-way ANOVA for groups of data. Because multiple comparisons were made for pairs of conditions and their statistical significance was estimated, the Bonferroni correction (Bland and Altman 1995) procedure was used, when appropriate, to adjust the threshold for the *p* value for statistical significance, which is usually set at 0.05 without this correction.

3. Results

3.1. Lung nodule volume reproducibility

All nodule-mimicking spheres were successfully segmented as judged by visual inspection. Tables 2 and 3 list the average volume errors calculated from the 3DAC boundary using equation (2) for the 4.8 mm diameter spheres from experiments I and II, respectively. The same imaging conditions were used in the two experiments conducted five months apart. The volume errors in these tables were analysed separately for each density (either 50 or 100 mg cc^{-1} of $CaCO_3$) and each scan. Six averages for each imaging condition are listed for each experiment, since there were three scans taken for each condition and two densities for each.

Table 2. Experiment I: the means and standard deviation of volume errors of 4.8 mm phantom nodules in each of the three repeated scans in the reproducibility study for various scanning and reconstruction parameters. For each set of parameters, there were five spheres with 50 mg cc^{-1} $CaCO_3$ and six spheres with 100 mg cc^{-1} $CaCO_3$ in each scan.

							Scan 1		Scan 2		Scan 3	
Pitch	Thickness (mm)	Interval (mm)	kVp	mA	Time (s)	Density (mg cc^{-1})	Average (%)	Standard deviation (%)	Average (%)	Standard deviation (%)	Average (%)	Standard deviation (%)
0.531:1	0.625	0.625	120	400	0.8	50	24.9	5.9	24.9	3.4	22.2	4.0
						100	19.2	5.0	20.4	6.7	15.7	6.0
1.375:1	1.25	1.25	120	80	0.5	50	24.4	7.3	29.0	7.0	25.5	4.5
						100	20.0	7.9	23.7	6.5	21.6	8.9
1.375:1	2.50	2.00	120	80	0.5	50	35.4	22.9	45.4	13.3	51.4	13.5
						100	43.5	9.1	48.1	4.3	44.1	7.8
1.375:1	0.625	0.625	140	300	0.5	50	21.3	4.7	21.8	4.9	20.2	6.7
						100	19.2	7.5	16.4	6.8	18.5	5.8
1.375:1	1.250	0.625	140	300	0.5	50	18.2	2.5	15.7	3.5	19.0	4.2
						100	17.0	5.9	20.0	5.0	16.0	5.7

Table 3. Experiment II (five months after experiment I): the means and standard deviation of volume errors for 4.8 mm phantom nodules in each of the three repeated scans in the reproducibility study for various scanning and reconstruction parameters. For each set of parameters, there were five spheres with 50 mg cc^{-1} $CaCO_3$ and five spheres with 100 mg cc^{-1} $CaCO_3$ in each scan.

							Scan 1		Scan 2		Scan 3	
Pitch	Thickness (mm)	Interval (mm)	kVp	mA	Time (s)	Density (mg cc^{-1})	Average (%)	Standard deviation (%)	Average (%)	Standard deviation (%)	Average (%)	Standard deviation (%)
0.531:1	0.625	0.625	120	400	0.8	50	21.3	6.0	21.2	4.7	21.5	4.6
						100	21.7	4.0	22.2	3.1	22.6	5.7
1.375:1	1.25	1.25	120	80	0.5	50	25.3	3.3	24.9	7.7	28.4	6.8
						100	22.8	5.9	21.7	5.5	25.0	7.5
1.375:1	2.50	2.00	120	80	0.5	50	50.4	10.2	48.7	11.8	50.2	9.8
						100	48.7	5.5	43.5	9.6	42.4	6.7
1.375:1	0.625	0.625	140	300	0.5	50	23.0	3.9	20.0	2.8	24.0	4.2
						100	21.5	6.3	21.3	2.6	19.2	6.7
1.375:1	1.250	0.625	140	300	0.5	50	21.4	6.2	20.8	4.2	17.6	3.4
						100	14.8	3.1	15.5	4.8	14.9	1.9

3.1.1. Analysis by density. The two-tailed paired *t*-test was performed where each pair consisted of the average volume error of the 50 mg cc^{-1} and the 100 mg cc^{-1} densities in one scan. There were five different imaging conditions in both experiments I and II, each of which had three identical scans, for a total of 30 pairs for statistical analysis. The volume error difference between the densities was not statistically significant ($p = 0.845$).

3.1.2. Reproducibility analysis. To determine whether there were differences in the volume errors among the three identical scans acquired on the same day, analysis of variance (ANOVA) was performed on the three identical scans. Each scan contained 11 (experiment I) or 10 (experiment II) nodules for each of the five different imaging conditions in experiments I and II. For a given imaging condition, the difference in average volume errors was not statistically significant ($p > 0.05$).

To determine reproducibility for scans separated by a few months, ANOVA was used to analyse the volume errors from the six scans (three from each experiment) taken under identical imaging conditions, each of which contained 11 or 10 nodules. For each of the five imaging conditions in experiments I and II, the difference in the volume measurements ($p > 0.05$) between the scans acquired five months apart did not achieve statistical significance.

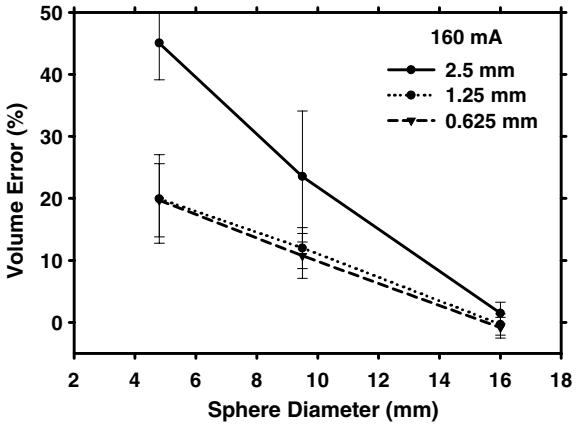


Figure 2. The effect of slice thickness on volume error at 120 kVp tube voltage, 160 mA tube current, 0.5 s rotation time, 0.625 mm slice interval, 1.375:1 pitch and 36 cm FOV. The error bars indicate ±1 SD of a measurement.

Table 4. Comparison of the volume errors for 4.8 mm diameter phantom nodules for various parameters taken during experiments I and II. The volume errors for the nodules with 50 mg cc^{-1} $CaCO_3$ and 100 mg cc^{-1} $CaCO_3$ are separately analysed. For each set of parameters, the averages and standard deviations are calculated for volume errors from three identical scans.

							Experiment I		Experiment II	
Pitch	Slice thickness (mm)	Slice interval (mm)	kVp	mA	Time (s)	Density (mg cc^{-1})	Average (%)	Standard deviation (%)	Average (%)	Standard deviation (%)
0.531:1	0.625	0.625	120	400	0.8	50	18.5	5.9	21.3	4.8
						100	24.0	4.4	22.2	4.8
1.375:1	1.25	1.25	120	80	0.5	50	21.8	7.5	26.2	6.0
						100	26.3	6.2	23.2	6.0
1.375:1	2.50	2.00	120	80	0.5	50	45.2	7.2	49.8	9.9
						100	44.1	17.3	44.9	7.5
1.375:1	0.625	0.625	140	300	0.5	50	18.0	6.5	22.3	3.8
						100	21.1	5.2	20.7	5.2
1.375:1	1.250	0.625	140	300	0.5	50	17.7	5.5	19.9	4.7
						100	17.6	3.6	15.1	3.2

3.1.3. Volume error variability. Table 4 shows a comparison of the volume errors averaged over all three scans for a given density and imaging condition. Higher variability of volume errors occurred with larger slice thicknesses and slice intervals. The greatest volume errors occurred for the scan using 2.5 mm slice thickness and 2.0 mm slice interval, which were the lowest resolution scans in this reproducibility study.

3.2. Effects of CT parameters on volume errors

3.2.1. Dependence on slice thickness. The average volume errors for varying slice thicknesses of 0.625 mm, 1.25 mm and 2.5 mm were compared for tube currents ranging from 80 to 320 mA. The other parameters were fixed at 1.375:1 pitch, 0.625 mm slice interval, 120 kVp tube voltage, 0.5 s rotation time and 36 cm FOV. The dependence of the average volume errors on nodule size is shown in figure 2 for a tube current of 160 mA as an example.

Table 5. Analysis of statistical significance in changes in average volume errors for varying nodule size, tube current and slice thickness values. Significant (with Bonferroni correction) and insignificant changes are marked with 'S' and 'NS', respectively. Other parameters were fixed at 1.375:1 pitch, 0.625 mm slice interval, 120 kVp tube voltage, 0.5 s rotation time and 36 cm FOV. The scan of 4.8 mm spheres at 80 mA and 1.25 mm slice thickness was not performed due to an oversight.

Nodule size (mm diameter)	Tube current (mA)	Slice thickness change (mm)		
		0.625 to 1.25	1.25 to 2.5	0.625 to 2.5
	80	N/A	N/A	S
4.8	160	NS	S	S
	320	NS	S	S
	80	NS	S	S
9.5	160	NS	NS	S
	320	NS	S	S
	80	NS	NS	S
16	160	NS	S	S
	320	NS	S	S

The trends are similar within the tube current range of 80–320 mA studied. The average volume errors of the 4.8, 9.5 and 16 mm spheres for all slice thicknesses and tube currents ranged from 19.4% to 45%, 9.9% to 23.5% and −0.8% to 1.7%, respectively.

For the 160 mA case, there were three different slice thickness values and three nodule sizes, resulting in nine paired t-tests. A p-value of less than $(0.05/9 = 0.0056)$ was considered statistically significant after application of the Bonferroni correction. There was no statistically significant difference in the average volume error ($p > 0.0056$) when the slice thickness changed from 0.625 to 1.25 mm for all nodule sizes. However, a change in slice thickness from 1.25 mm to 2.5 mm resulted in a statistically significant ($p < 0.0056$) difference in the average volume error for the 4.8 mm and 16 mm nodules, but not for the 9.5 mm nodules ($p = 0.0073$). A change in slice thickness from 0.625 mm to 2.5 mm affected the average volume error significantly ($p < 0.0056$) for all nodule sizes. The results for the various nodule sizes and all three tube currents are summarized in table 5.

3.2.2. Dependence on tube current (mA). For each slice thickness, three different tube currents of 100, 200 and 400 mA were used to scan the three different nodule sizes. Table 6 lists the standard deviations of the HU values in several regions of interest of a slice intersecting approximately the centre of the nodules in each scan. The standard deviations represented the relative noise levels in the CT scans due to the tube current changes. As expected, the noise decreased as the slice thickness increased and the tube current decreased. It was found that changing the noise level within the range studied did not significantly ($p > 0.05$) affect the average volume error of the nodules for all nodule sizes.

For the 4.8 mm spheres, the volume errors for varying slice thickness and tube current are shown in figure 3. The parameters fixed were 0.531:1 pitch, 0.625 mm slice interval, 120 kVp, 0.8 s rotation time and 36 cm FOV. Note that the pitch and rotation time for this set of parameters differ from those used for the data shown in figure 2 and table 5, although the trends are similar. With a slice thickness of 0.625 mm, the average volume errors were 21.6%, 20.8% and 21.7% for 100, 200 and 400 mA scans, respectively. When the slice thickness was increased to 2.5 mm, the corresponding errors increased to 42.3%, 43.3% and 40.7% for the three tube currents, respectively.

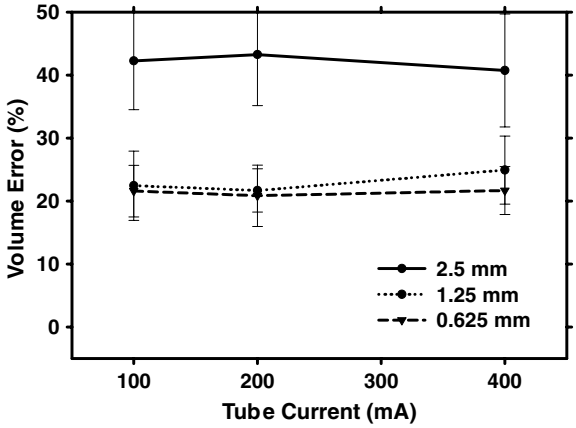


Figure 3. Effect of slice thickness and tube current on volume error for 4.8 mm spheres. The parameters fixed were 0.531:1 pitch, 0.625 mm slice interval, 120 kVp, 0.8 s rotation time and 36 cm FOV. The error bars indicate ± 1 SD of a measurement.

Table 6. Standard deviations of HU values of regions of interest (ROIs) on CT slices. The ROIs were placed in the heart muscle, lung parenchyma and muscle near the vertebra. Since the chest phantom was not moved when the scanning parameters were changed, the slice and position of the ROIs were the same for all entries of this table. The parameters kept constant were 0.531:1 pitch, 0.625 mm slice interval, 120 kVp, 0.8 s rotation time and 36 cm FOV.

Slice thickness (mm)	Tube current (mA)	Heart	Parenchyma	Muscle
	100	11.78	9.29	16.84
0.625	200	8.60	6.66	11.16
	400	6.38	5.07	10.98
	100	10.02	8.11	15.03
1.25	200	7.73	6.04	11.12
	400	5.92	4.86	9.51
	100	7.39	5.78	9.25
2.5	200	5.48	4.61	9.07
	400	4.50	3.58	8.57

There was no statistically significant change (all $p > 0.6$) in average volume errors when slice thickness changed from 0.625 to 1.25 mm for the 100, 200 and 400 mA scans. However, further increase in slice thickness from 0.625 mm or 1.25 mm to 2.5 mm resulted in a statistically significant ($p < 0.0056$) difference in average volume error for all tube currents.

3.2.3. Dependence on pitch. The effects of varying pitch from 0.531:1, 0.969:1 and 1.375:1 are shown in figure 4. The CT parameters were fixed at 0.625 mm slice thickness and slice interval, 120 kVp tube voltage, 400 mA tube current, 0.8 s rotation time and 36 cm FOV. The volume error decreased as the nodule size increased, similar to the trends observed in figure 2. The average volume errors of the 4.8, 9.5 and 16 mm spheres ranged from 20.6% to 21.7%, 10.4% to 10.6% and −0.4% to −0.6%, respectively. The variation in pitch did not significantly affect the volume error of segmentation ($p > 0.05$) within the range studied for all nodule sizes.

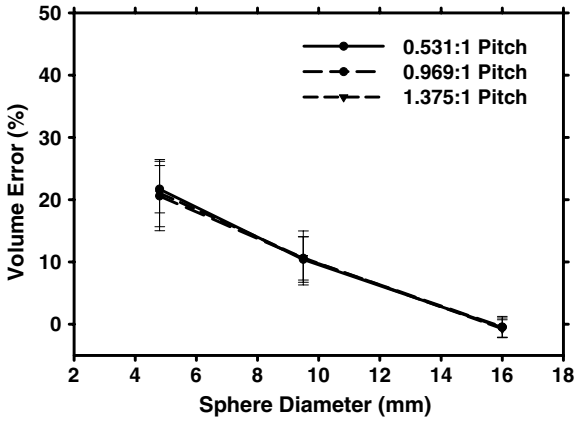


Figure 4. Effect of pitch on volume error. The scanning parameters were fixed at 0.625 mm slice thickness and slice interval, 120 kVp tube voltage, 400 mA tube current, 0.8 s rotation time and 36 cm FOV. The error bars indicate ± 1 SD of a measurement.

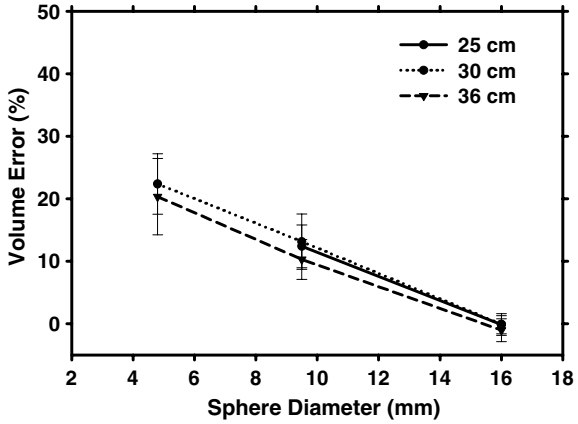


Figure 5. Effect of the scanning field of view (FOV) on volume error. The other parameters were fixed at 0.531:1 pitch, 0.625 mm slice thickness and interval, 120 kVp tube voltage, 400 mA tube current and 0.8 s rotation time. The error bars indicate ±1 SD of a measurement.

3.2.4. Dependence on field of view. The effects of varying FOV values from 25 cm to 36 cm on volume error are presented in figure 5. Note that increasing the FOV increases the pixel size on the axial plane and also may affect the interpolated voxel size used for the segmentation. The other parameters were fixed at 0.531:1 pitch, 0.625 mm slice thickness and interval, 120 kVp tube voltage, 400 mA tube current and 0.8 s rotation time. The CT scan containing 4.8 mm spheres with 25 cm FOV were not used because the beginning and end slice acquisition location were set erroneously, resulting in spheres that were not completely scanned. The average volume errors of the 4.8, 9.5 and 16 mm spheres ranged from 20.3% to 22.4%, 10.3 to 12.4% and −1.0% to −0.1% respectively. There was no statistically significant change ($p > 0.006$, Bonferroni correction) in the average volume errors within the range of FOV studied.

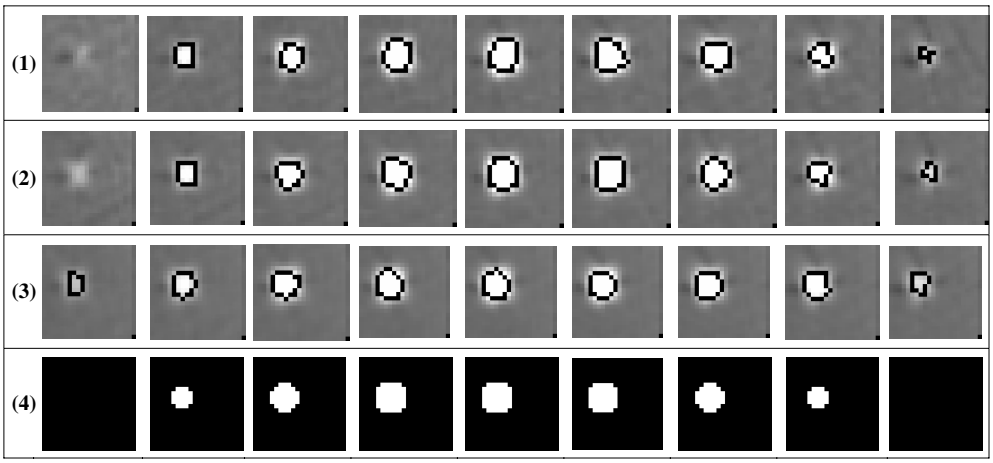


Figure 6. Images of all slices with 3DAC segmented contours of a 4.8 mm phantom nodule and various slice thicknesses. Row (1), 0.625 mm, 17% error. Row (2), 1.25 mm, 22% error. Row (3), 2.5 mm, 24% error. Other parameters were kept constant: 0.625 mm slice interval, 120 kVp, 320 mA, 0.5 s rotation time, 36 cm FOV. Note the additional slice at 2.5 mm slice thickness that was segmented due to the high voxel intensity caused by partial volume averaging. Row (4) contains a computer-generated sphere of 4.8 mm diameter, symmetrically aligned with the pixel array, at 0.625 mm slice interval for comparison.

3.3. Partial volume effects on volume errors

The overestimation in nodule volume is mainly caused by partial volume effects. Typical segmented contours for the 4.8 mm nodule at three slice thicknesses and the same 0.625 mm slice interval are shown in figure 6, rows (1)–(3). The segmented boundaries are visually reasonable although the average volume errors of 20%–45% (see figure 3) seem excessive. The segmented volume increases as slice thickness increases although the true nodule size is the same. To further demonstrate the effect, a spherical object of radius 4.8 mm was digitally generated and is shown in figure 6, row (4). Comparison of both the digitally generated slices and the actual CT slices showed that partial volume effects generally cause the volume to appear larger than the true volume. This is because the nodule boundary is blurred and more pixels that are outside the true nodule boundary become brighter and are considered part of the nodule. There are also additional slices that appear to contain part of the nodule while the true nodule may not have intersected those slices. These effects become stronger when the slice thickness increases.

3.4. Manual segmentation by a radiologist

Figure 7 shows a comparison of a computer-generated boundary of a discretized 4.8 mm sphere, the 3DAC result, and the radiologist's hand-drawn boundaries. The average volume errors of the manually segmented volumes for the five 4.8 mm diameter nodules in each of the 0.625, 1.25 and 2.5 mm slice thicknesses were 239.3 $\pm$ 36.5%, 214.8 $\pm$ 34.4% and 275.6 $\pm$ 40.3%, respectively. The corresponding 3DAC average volume errors for the same nodules and imaging conditions were 18.6 $\pm$ 5.5%, 23.6 $\pm$ 7.6% and 45.2 $\pm$ 5.4%, respectively.

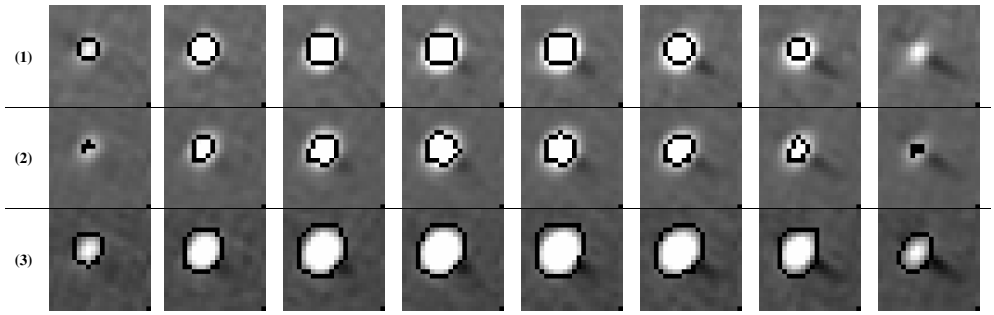


Figure 7. Comparisons of contours for a 4.8 mm diameter nodule. Imaging conditions were 1.375:1 pitch, 0.625 mm slice thickness and interval, 120 kVp, 160 mA, 0.5 s rotation and 36 cm FOV. Row (1), contours of a computer-generated 4.8 mm diameter discretized sphere; row (2), 3DAC output; row (3), radiologist segmentation.

4. Discussion

In this paper, our focus is to evaluate the effects of CT scanning and reconstruction parameters on the accuracy and reproducibility of automated nodule volume estimation. Although the absolute magnitude of the volume errors estimated using phantom nodules may be different from those of real nodules in patient scans, this study reveals important trends for the dependence of the volume errors on CT imaging conditions and the variability of these measurements. This information may serve as a guide when automated or manual methods are used to assess interval volume change of a nodule from serial CT scans.

In one part of this study, we asked a fellowship-trained thoracic radiologist to segment nodules imaged with varying slice thicknesses. The radiologist's hand-drawn boundaries appear to have included the partial volume voxels. Quantitatively, the volume errors of the manual segmentation ranged from 150% to 350%, compared to the 9% to 52% for the 3DAC. Radiologists are not required to outline lesions in their clinical practice. They are trained to estimate whether a nodule changes size in serial CT scans. The judgment of where the lesion boundaries are on an image is subjective, as indicated by the large inter- and intra-observer variabilities among experienced chest radiologists in the LIDC study (Meyer *et al* 2006). Since the percentage volume change for a given diameter change depends strongly on the nodule size, an overestimation of the nodule diameter can cause large underestimation in nodule growth rate for small nodules. This underscores the difficulty that partial volume effects will impose on the assessment of nodule growth. The advantages of automated volume segmentation over manual segmentation include immunity from variability due to changing window and level settings. Radiologists may change the settings to better view the nodule, resulting in differing perceptions of the boundary and thus different volumes (Harris *et al* 1993). Computerized analysis uses the CT numbers of the voxels themselves in a deterministic algorithm and it will produce exactly the same result for identical input. However, even for CT scans under identical imaging conditions, the acquired images contain statistical variations in the x-ray photons recorded at the detector and other uncertainties of the CT scanner. For example, the starting scan position of the CT scanner is not perfectly reproducible. The slice locations relative to the anatomical structures are therefore not identical in repeated scans even if the phantom is not repositioned.

The reproducibility of volume measurements from scans taken on the same day or five months apart were evaluated in this study. The comparison of segmented volume errors

for CT scans with identical imaging conditions between the two experiments shows that, on average, the volume errors agree to within a few per cent in the absence of motion or other physiological changes associated with a real patient. This consistency was also observed by Wormanns (Wormanns *et al* 2004), who reported high precision in the lung nodule volume measurement based on an automated method applied to two scans taken 10 min apart using the same imaging conditions.

Our study indicated that the average volume error relative to the ground truth was consistently overestimated, except for the 16 mm nodules at a thin slice thickness that occasionally showed slight underestimation within experimental uncertainties. The overestimation is mainly caused by partial volume averaging rather than the AC segmentation settings as demonstrated in our analysis of the partial volume effect. The nodule size and slice thickness had the largest effects on the measured volume accuracy. For a given CT condition, the smaller the nodules, the higher the percentage volume errors.

We chose the nodule sizes used in this study because there is general consensus that pulmonary nodules of clinical significance in CT evaluation have longest diameters in the range of 3–30 mm. Nodules less than 10 mm should be followed up with CT (Ost *et al* 2003, Henschke *et al* 1999), and we chose the 4.8 mm and 9.5 mm to represent the low and high end of the range. In addition, 15 to 22 mm nodules are at intermediate risk for cancer, so we chose 16 mm as the representative size. Our results indicate that, for nodules greater than 16 mm, the volume errors due to the factors considered in the current study are negligible so that the size range of the nodules of interest has been covered.

In this study, typical CT scans have 0.703 mm resolution in the axial plane. The voxels are isotropic after interpolation to a volume of $(0.703)^3$ mm^3. At this voxel size, a 4.8 mm diameter sphere contains approximately 168 voxels if the effect of discretization is ignored. Seven slices (4.8 mm/0.703 mm) is the smallest number of slices that can represent this nodule. A 20% over-estimation in volume is the result of 33 extra voxels, which corresponds to less than five pixels per slice. In other words, if five more voxels in each slice are considered part of the nodule instead of the background for each slice, it will result in a 20% volume error. Thus, the volume error for small nodules is especially sensitive to the uncertainties in the segmented boundary, as a slight deviation due to partial volume effect and reconstruction artefacts such as a blurry or irregular edge would result in substantial percentage error. When the slice thickness is large, the blurred boundary due to partial volume averaging contributes to extra slices for the nodule. The additional slices and additional voxels in each slice together would have caused the high volume errors ($>$40% for 2.5 mm slice thicknesses) for the 4.8 mm nodules.

This is contrasted to a study by Yankelevitz *et al* that reported volume errors of $\pm$3% for 3.2 and 3.96 mm diameter phantom nodules (Yankelevitz *et al* 2000). The difference may be attributed to two main factors. First, Yankelevitz *et al* scanned the phantom at a high resolution of 1 mm beam collimation and 9.6 cm FOV, resulting in a pixel size of 0.188 $\times$ 0.188 mm in the axial plane, with a reconstructed slice interval of 0.5 mm and trilinear interpolation to obtain isotropic voxels. In our study, the smallest slice interval of 0.625 mm would have resulted in interpolated isotropic voxels with volumes of 0.244 mm^3, which is 36.7 times larger than their isotropic voxels with volumes of 0.006 65 mm^3 ($= 0.188^3$). Second, our AC parameters were trained with patient nodules (Way *et al* 2006) rather than the current set of phantom nodules. We can expect that if the AC parameters were trained to fit the edge characteristics of phantom nodules, the AC segmented volumes could be adjusted to be closer to their ground-truth volumes. However, the resulting AC algorithm may not perform well for patient nodules. Because segmentation of patient nodules is the goal for developing

the automated AC method, we chose to apply the trained AC to phantom nodules without retraining.

The CT scanning pitch did not significantly affect the volume error ($p > 0.05$) for phantom nodules of all sizes when the other parameters were fixed as chosen in this study. However, this experiment was only performed for thin slices (0.625 mm thickness and interval). Future experiments are needed to determine whether this result would still hold with larger slice thicknesses and intervals. There was also no statistically significant effect of FOV on the average volume error, although the average volume error for the 36 cm FOV was consistently slightly lower than that for the 25 cm or 30 cm FOV. One may expect larger errors for larger FOVs because of the increase in the voxel size and thus the partial volume blurring. However, a larger voxel size could result in a smaller segmented volume due to the poorer approximation to a sphere by the large discrete voxels. For simplicity, consider an example in 2D. Because the segmented nodule contour was represented by a polygon, smaller voxels would allow more vertices to form the polygon. Consider a polygon inscribed in a circle by placing five vertices on the border of the circle. The area of this polygon would be smaller than that if ten vertices were placed on the border of the circle. Since the segmented nodule volume was generally over-estimated, the reduced volume due to increased voxel size could reduce the volume error if the reduction was greater than the increase due to the increased partial volume effect.

There was no statistically significant dependence of volume errors on tube current for pitch settings of 0.531:1 and 1.375:1 and for slice thicknesses from 0.625 mm to 2.5 mm. The volume errors therefore were relatively independent of dose for the phantom nodules. For assessment of nodule growth, it may be sufficient to use high resolution but low-dose serial scans.

We have demonstrated that there is a large difference in the estimated volume when the CT scan is acquired with different imaging conditions. To minimize the error, thin slice CT scans should be used, especially for small nodules. Furthermore, to estimate the growth of a nodule, it is important to use the same scanning and reconstruction parameters in follow-up scans. Currently, follow-up scanning is performed for nodules after certain time intervals, but there are no specifications on the parameters to be used. A baseline scan and follow-up scans using different slice thicknesses, for example, may result in similar volume measurements when in fact the nodule size has changed. The error in volume change estimation may lead to misdiagnosis and delay in treatment.

The volume errors also raise questions regarding protocols used to screen for lung cancer. Protocols such as the NLST trial that used 2.5 mm slice thickness and 2.0 mm slice intervals may be too thick to accurately assess volume and growth. A thin-slice protocol should be used or a work-up scan at a higher resolution should be performed after a nodule is detected. However, this needs to be balanced with other considerations for screening, such as the larger number of images to be interpreted and archived associated with thin-slice CT screening, patient dose and the costs of work-up.

In clinical CT scans, variabilities such as patient motion and the change in nodule size and boundary characteristics over time will further degrade the reproducibility between serial exams. Furthermore, clinical nodules in general have less sharp and more irregular boundaries than spherical phantom nodules. These characteristics may cause additional volume errors compared with those estimated in this study, and cause large inter- and intra-observer variabilities even for experienced chest radiologists (Meyer *et al* 2006). These errors will be impossible to estimate because there is no ground truth volume for clinical nodules and the cause and magnitude of the errors may change from case to case. With computer segmentation, although the percentage volume errors for small nodules are large,

the reproducibility of volume estimation from repeated scans is within a few per cent. It may be expected that if a consistent CT protocol and computerized segmentation algorithm are used for serial CT scans to assess nodule growth, the error in assessing the volume change could be less than the absolute volume error.

There are limitations in this study. First, it will be of strong interest to estimate the smallest possible volume change that can be estimated with confidence using automated segmentation. Second, it is not known whether the same trends observed in this phantom study would be seen for real lung nodules. Third, it is also not known whether the dependence of volume errors on imaging and reconstruction parameters is consistent for CT scans acquired with scanners from different manufacturers. Fourth, for the evaluation of CT parameters on volume errors, we fixed the slice interval at 0.625 mm to reduce the number of variables. The effects of this parameter and its interaction with other parameters on volume error are therefore still unknown. Finally, we did not employ targeted reconstructions with smaller FOVs such as 9.6 cm that would reduce volume averaging in the axial plane. These and other issues will be investigated in future studies.

In summary, we have found that scanning and reconstruction parameters of CT scans affect automatic volume measurement by the 3DAC method. This investigation has important clinical implications because comparing nodule volumes measured from two different scans to determine whether there is growth is a commonly used method in initial diagnosis of lung cancers. In the larger context of a computerized image analysis system, this shows that not only is the segmentation algorithm important, but the method of image acquisition for the CT scans used in the segmentation also affects the outcome. Thus, to accurately follow up on a nodule to detect interval change in volume, the scanning and reconstruction parameters should be properly chosen and kept constant between the initial and follow-up scans to minimize the variability in the volume change evaluation.

Acknowledgments

This work is supported by USPHS grant No. CA 93517. The authors thank Emmanuel G Christodoulou, PhD, and Sandra C Larson, PhD, for assisting with the CT image acquisition.

References

Armato S G, Altman M B and Wilkie J 2003 Automated lung nodule classification following automated nodule detection on CT: a serial approach *Med. Phys.* **30** 1188–97

Bland J M and Altman D G 1995 Statistics notes: multiple significance tests: the Bonferroni method *Br. Med. J.* **310** 170

Bogot N R, Kazerooni E A, Kelly A M, Quint L E, Desjardins B and Nan B 2005 Interobserver and intraobserver variability in the assessment of pulmonary nodule size on CT using film and computer display methods *Acad. Radiol.* **12** 948–56

Diederich S, Wormanns D, Semik M, Thomas M, Lenzen H, Roos N and Heindel W 2002 Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers *Radiology* **222** 773–81

Erasmus J, Gladish G, Broemeling L, Sabloff B, Truong M, Herbst R and Munden R 2003 Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response *J. Clin. Oncol.* **21** 2574–82

Fischbach F, Knollmann F, Griesshaber V, Freund T, Akkol E and Felix R 2003 Detection of pulmonary nodules by multislice computed tomography: improved detection rate with reduced slice thickness *Eur. Radiol.* **13** 2378–83

Ge Z, Sahiner B, Chan H P, Hadjiiski L M, Cascade P N, Bogot N, Kazerooni E A, Wei J and Zhou C 2005 Computer aided detection of lung nodules: false positive reduction using a 3D gradient field method and 3D ellipsoid fitting *Med. Phys.* **32** 2443–54

Goo J M, Tongdee T, Tongdee R, Yeo K, Hildebolt C F and Bae K T 2005 Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy *Radiology* **235** 850–6

Goodsitt M M, Chan H-P, Way T W, Larson S C, Christodoulou E G and Kim J 2006 Accuracy of the CT numbers of simulated lung nodules imaged with multi-detector CT scanners *Med. Phys.* **33** 3006–17

Gurcan M N, Sahiner B, Petrick N, Chan H P, Kazerooni E A, Cascade P N and Hadjiiski L 2002 Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system *Med. Phys.* **29** 2552–8

Harris K, Adams H, Lloyd D and Harvey D 1993 The effect on apparent size of simulated pulmonary nodules of using three standard CT window settings *Clin. Radiol.* **47** 241–4

Henschke C I *et al* 1999 Early lung cancer action project: overall design and findings from baseline screening *The Lancet* **354** 99–105

Im J, Gamsu G, Gordon D, Stein M, Webb W, Cann C and Niklason L 1988 CT densitometry of pulmonary nodules in a frozen human thorax *Am. J. Roentgenol.* **150** 61–6

Kaneko M, Eguchi K, Ohmatsu H, Kakinuma R, Naruke T, Suemasu K and Moriyama N 1996 Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography *Radiology* **201** 798–802

Kass M, Witkin A and Terzopoulos D 1987 Snakes: active contour models *Int. J. Comput. Vis.* **1** 321–31

Kawata Y, Niki N, Ohmatsu H, Kakinuma R, Eguchi K, Kaneko M and Moriyama N 1998 Quantitative surface characterization of pulmonary nodules based on thin-section CT images *IEEE Trans. Nucl. Sci.* **45** 2132–8

Ko J P, Rusinek H, Jacobs E L, Babb J S, Betke M, McGuinness G and Naidich D P 2003 Small pulmonary nodules: volume measurement at chest CT—phantom study *Radiology* **223** 864–70

Li F, Aoyama M, Shiraishi J, Abe H, Li Q, Suzuki K, Engelmann R, Sone S, MacMahon H and Doi a K 2004 Radiologists' performance for differentiating benign from malignant lung nodules on high-resolution CT using computer-estimated likelihood of malignancy *Am. J. Roentgenol.* **183** 1209–15

Lillington G A and Caskey C I 1993 Evaluation and management of solitary and multiple pulmonary nodules *Clin. Chest Med.* **14** 111–9

MacMahon H, Austin J H M, Gamsu G, Herold C J, Jett J R, Naidich D P, Patz E F Jr and Swensen S J 2005 Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society *Radiology* **237** 395–400

Meyer C R, Johnson T D, McLennan G, Aberle D R, Kazerooni E A, MacMahon H, Mullan B F, Yankelevitz D F, van Beek E J R and Armato S G III 2006 Evaluation of lung MDCT nodule annotation across radiologists and methods *Acad. Radiol.* **13** 1254–65

Nawa T, Nakagawa T, Kusano S, Kawasaki Y, Sugawara Y and Nakata H 2002 Lung cancer screening using low-dose spiral CT: results of baseline and 1-Year follow-up studies *Chest* **122** 15–20

Ost D, Fein A M and Feinsilver S H 2003 The solitary pulmonary nodule *N. Engl. J. Med.* **348** 2535–42

Petkovska I, Brown M S, Goldin J G, Kim H J, McNitt-Gray M F, Abtin F G, Ghurabi R J and Aberle D R 2007 The effect of lung volume on nodule size on CT *Acad. Radiol.* **14** 476–85

Shah S K, McNitt-Gray M F, Rogers S R, Goldin J G, Suh R D, Sayre J W, Petkovska I, Kim H J and Aberle D R 2005 Computer aided characterization of the solitary pulmonary nodule using volumetric and contrast enhancement features *Acad. Radiol.* **12** 1310–9

Sone S *et al* 2001 Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner *Br. J. Cancer* **84** 25–32

Sone S, Takashima S, Li F, Yang F, Honda T, Maruyama Y, Hasega M, Yamanda T, Kubo K, Hanamura K and Asakura K 1998 Mass screening for lung cancer with mobile spiral computed tomography scanner *The Lancet* **352** 1242–5

Swensen S J, Jett J R, Hartman T E, Midthun D E, Mandrekar S J, Hillman S L, Sykes A-M, Aughenbaugh G L and Allen A O B L 2005 CT screening for lung cancer: five-year prospective experience *Radiology* **235** 259–65

Swensen S J, Jett J R, Hartman T E, Midthun D E, Sloan J A, Sykes A M, Aughenbaugh G L and Clemens M A 2003 Lung cancer screening with CT: Mayo Clinic experience *Radiology* **226** 756–61

Tan B B, Flaherty K R, Kazerooni E A and Iannettoni M D 2003 The solitary pulmonary nodule *Chest* **123** 89–96

Usuda K, Saito Y, Sagawa M, Sato M, Kanma K, Takahashi S, Endo C, Chen Y, Sakurada A and Fujimura S 1994 Tumor doubling time and prognostic assessment of patients with primary lung cancer *Cancer* **74** 2239–44

Way T W, Hadjiiski L M, Sahiner B, Chan H-P, Cascade P N, Kazerooni E A, Bogot N and Zhou C 2006 Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours *Med. Phys.* **33** 2323–37

Williams D J and Shah M 1992 A fast algorithm for active contours and curvature estimation *CVGIP: Image Underst.* **55** 14–26

Winer-Muram H T, Jennings S G, Meyer C A, Liang Y, Aisen A M, Tarver R D and McGarry R C 2003 Effect of varying CT section width on volumetric measurement of lung tumors and application of compensatory equations *Radiology* **229** 184–94

Wormanns D, Kohl G, Klotz E, Marheine A, Beyer F, Heindel W and Diederich S 2004 Volumetric measurements of pulmonary nodules at multi-row detector CT: *in vivo* reproducibility *Eur. Radiol.* **14** 86–92

Yankelevitz D F, Gupta R, Zhao B and Henschke C I 1999 Small pulmonary nodules: evaluation with repeat CT —preliminary experience *Radiology* **212** 561–6

Yankelevitz D, Reeves A, Kostis W, Zhao B and Henschke C 2000 Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation *Radiology* **217** 251–6

Zerhouni E A, Spivey J F, Morgan R H, Leo F P, Stitik F P and Siegelman S S 1982 Factors influencing quantitative CT measurements of solitary pulmonary nodules *J. Comput. Assist. Tomogr.* **6** 1075–87