Accuracy of registration algorithms in subtraction CT of the lungs: A digital phantom study

Dagmar Grob and Luuk Oostveen
Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands

Jan Rühaak and Stefan Heldmann
Fraunhofer Institute for Medical Image Computing MEVIS, Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Maria-Goeppert-Str. 3, 23562 Lübeck, Germany

Brian Mohr
Canon Medical Research Europe, Anderson Place 2, E6 5NP Edinburgh, Scotland

Koen Michielsen
Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands

Sabrina Dorn
German Cancer Research Center, Heidelberg (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Mathias Prokop
Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands

Marc Kachelrieβ
German Cancer Research Center, Heidelberg (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Monique Brink
Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands

Ioannis Sechopoulos[a)]
Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands
Dutch Expert Center for Screening (LRCB), Wijchenseweg 101, 6538 SW Nijmegen, The Netherlands

(Received 19 January 2018; revised 15 February 2019; accepted for publication 7 March 2019; published 8 April 2019)

Purpose: The purpose of this study was to assess, using an anthropomorphic digital phantom, the accuracy of algorithms in registering precontrast and contrast-enhanced computed tomography (CT) chest images for generation of iodine maps of the pulmonary parenchyma via temporal subtraction.
Materials and methods: The XCAT phantom, with enhanced airway and pulmonary vessel structures, was used to simulate precontrast and contrast-enhanced chest images at various inspiration levels and added CT simulation for realistic system noise. Differences in diaphragm position were varied between 0 and 20 mm, with the maximum chosen to exceed the 95th percentile found in a dataset of 100 clinical subtraction CTs. In addition, the influence of whole body movement, degree of iodine enhancement, beam hardening artifacts, presence of nodules and perfusion defects in the pulmonary parenchyma, and variation in noise on the registration were also investigated. Registration was performed using three lung registration algorithms — a commercial (algorithm A) and a prototype (algorithm B) version from Canon Medical Systems and an algorithm from the MEVIS Fraunhofer institute (algorithm C). For each algorithm, we calculated the voxel-by-voxel difference between the true deformation and the algorithm-estimated deformation in the lungs.
Results: The median absolute residual error for all three algorithms was smaller than the voxel size ($1.0 \times 1.0 \times 1.0$ mm^3) for up to an 8 mm diaphragm difference, which is the average difference in diaphragm levels found clinically, and increased with increasing difference in diaphragm position. At 20 mm diaphragm displacement, the median absolute residual error after registration was 0.85 mm (interquartile range, 0.51–1.47 mm) for algorithm A, 0.82 mm (0.50–1.40 mm) for algorithm B, and 0.91 mm (0.54–1.52 mm) for algorithm C. The largest errors were seen in the paracardiac regions and close to the diaphragm. The impact of all other evaluated conditions on the residual error varied, resulting in an increase in the median residual error lower than 0.1 mm for all algorithms, except in the case of whole body displacements for algorithm B, and with increased noise for algorithm C.

Conclusion: Motion correction software can compensate for respiratory and cardiac motion with a median residual error below 1 mm, which was smaller than the voxel size, with small differences among the tested registration algorithms for different conditions. Perfusion defects above 50 mm will be visible with the commercially available subtraction CT software, even in poorly registered areas, where the median residual error in that area was 7.7 mm. [https://doi.org/10.1002/mp.13496]

Key words: perfusion defect, registration algorithms, subtraction CT, thorax CT, voxel by voxel comparison

1. INTRODUCTION

Distribution of intravenously administered iodinated contrast agent in the lungs can be used as a marker for pulmonary perfusion, for instance for assessing severity of pulmonary vascular obstruction, compression, or constriction,[1,2] and for characterization of lung masses[3] and of interstitial lung disease.[4] Various computed tomography (CT) techniques exist to image the iodine distribution in the pulmonary parenchyma. One option is to perform subtraction CT (Fig. 1), which has been implemented in clinical practice to depict perfusion defects caused by pulmonary embolism.[5,6] This technique generates an iodine map by subtracting a precontrast CT image from a contrast-enhanced CT image. After removal of vessels, filtering, and adding an appropriate color scale, this iodine map reflects iodine distribution in the pulmonary parenchyma. However, an optimal temporal subtraction is only possible if location differences are absent and the two images are identical in geometry, so that any differences between the images are only caused by the iodinated contrast.

Currently, the accuracy of the registration algorithm, for example, whether or not it can compensate for the influence of respiratory and cardiac motion between two consecutive scans, which is the limited factor for subtraction CT imaging to be in widespread use,[7,8] is the major discussion point, and therefore, it is the focus of this paper. Some registration algorithms have been evaluated in challenges such as the EMPIRE10 image registration challenge[9] with resulting accuracies of the landmarks for the top ten algorithms between 0.63 and 0.84 mm. However, these evaluations focused on lung imaging for nodule characterization, and therefore only used a limited number of landmarks. In these studies, landmarks were manually selected, annotated, and matched to the closest observer choice. This explains why distance values varied from 0 mm (perfect agreement with one of the observers) up to 67.1 mm.[9] Second, potentially relevant locations for pathology depiction, such as parenchymatous pulmonary locations, do not contain landmarks and therefore cannot be evaluated in such a way. Overall, determining if a registration algorithm for subtraction CT is sufficiently accurate for a specific diagnostic application depends on the size scale of said application, as found for registration in radiotherapy applications.[10] For example, perfusion defects caused by most clinically relevant pulmonary embolisms are in centimeter range.[11] In addition, for the clinical task of assessing the lung parenchyma for perfusion imaging, evaluation of movement across all lung voxels is a better performance metric than using a limited number of landmarks. Therefore, the goal of this study was to evaluate the accuracy of motion correction under a variety of conditions in three registration algorithms for subtraction chest CT.

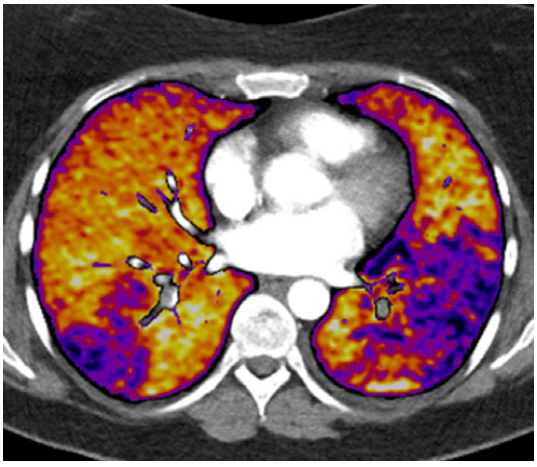

FIG. 1. 3 mm axial reconstruction of CT pulmonary angiography (CTPA) with an iodine map created with subtraction CT, with a 100 kV CTPA image as underlay. The image displays wedge-shaped perfusion defects caused by bilateral pulmonary emboli. [Color figure can be viewed at wileyonlinelibrary.com]

2. METHODS AND MATERIALS

2.A. Anthropomorphic phantom

To compare the voxel-by-voxel alignment of two images after registration, a digital phantom with known ground truth was used. These were generated using the improved 4D extended cardiac-torso (XCAT) anthropomorphic digital phantom (Duke University, Durham, NC, USA) at different inspiration levels and cardiac phases, in addition to other varying conditions like whole body displacement, different intravenous iodine concentration, presence or absence of beam hardening artifacts simulated by adding titanium pins in the spine, addition of nodules in the lungs, wedge-shaped perfusion defects in the pulmonary parenchyma of the contrast-enhanced image, and different noise level. This digital phantom provides a virtual model of patient anatomy and physiology including pulmonary vasculature and bronchopulmonary structures up to the terminal branches, without

collision of structures.[12,13] The phantom, based on ICRP Publication 89 data, is commonly used in imaging research due to the flexibility of its nonuniform rational basis spline (NURBS)-based surface primitives, variation in anatomy, motion, and variable spatial resolution, that can be simulated. The patient motion was modeled using 4D-tagged magnetic resonance imaging data (cardiac motion) and 4D high-resolution respiratory-gated CT data (respiratory motion).[12] These advantages also make it ideal for evaluation of registration algorithms.[13]

The XCAT phantom allows for the specification of more than 200 parameters, including contrast administration comparable to clinical practice, patient and scan characteristics, and respiratory and cardiac phase. This allows manipulation of the voxel values representing different tissues with enhanced attenuation due to the presence of intravenous-iodinated contrast agent. In this way, phantoms could be simulated representing a patient being scanned at different respiratory phases with and without contrast.

The following settings were used for the XCAT phantom realizations: inclusion of the entire chest, arm position above the shoulders, respiratory motion (12 breaths per minute) and cardiac motion (60 bpm), 12 output frames, isotropic voxels ($1.0 \times 1.0 \times 1.0$ mm^3), attenuation values of the spine were set equivalent to those of the ribs, making them more realistic, and 64 subvoxels/voxel. Subvoxels are used to estimate the mean attenuation of all the tissues that are contained within a single voxel with higher accuracy. All other parameters were set to default.

2.B. Variations in phantom

First, to study clinically relevant inspiration differences, 100 clinical subtraction CT pulmonary angiography cases from our institution were retrospectively evaluated by measuring the difference in diaphragm positions between precontrast and contrast-enhanced scans. Patients received instructions to hold their breath after shallow breathing during both scans. It was found that the mean difference was 6 mm, with 20 mm being the 98th percentile. One contrast-enhanced image was generated at the first moment of the cardiac and breathing phases, followed by 11 different output images with a variation in diaphragm level and cardiac phase at a respiratory cycle of 5 s (Table I, columns 1 and 2). Figure 2 shows examples of two images at different respiratory and cardiac phases (at 0% and 8.3% of the RR interval) with and without iodine contrast. To test for systematic registration errors, pairs of images were generated at two identical respiratory and cardiac phases: one containing intravenous contrast and one without intravenous contrast.

Second, to investigate whole body movement of the patient or movement caused by table movement, the images of the precontrast scan were shifted in the X, Y, and Z directions. A displacement of 30 mm was introduced in the X and Y directions and a 20 mm displacement in Z direction. In all situations, the entire lung was included in the CT image. The influence of these displacements was tested with identical cardiac and breathing phase in the precontrast and contrast-enhanced image and with 8 mm (average clinical difference and closest available to 6 mm) and 20 mm diaphragm difference (largest clinical difference) of the precontrast image compared to the contrast-enhanced image.

Third, to study the influence of iodine enhancement, the concentration of the iodine was varied by a factor of 2 and 0.5. The influence of varying enhancement level was again tested using the identical precontrast and contrast-enhanced phase, and with the 8 and 20 mm difference in diaphragm level between the precontrast and contrast-enhanced scans. The enhancement of iodine was determined with a circular region-of-interest (ROI) in the pulmonary truncus after CT simulation of the images was obtained.

Fourth, the influence of beam hardening artifacts was investigated. Since the artifacts from the ribs and spine were minor, three pairs of titanium pedicle screws with two connecting vertical rods were added to the thoracic spine (T6–T8). The screws and rods had a diameter of 8 mm. The registration algorithms were tested for the 0, 8, and 20 mm difference and the influence on the iodine map was shown.

Finally, two relevant pathological pulmonary entities were implemented in the phantom: nodules and perfusion defects

TABLE I. The median, interquartile range, and the 95th percentile of the true displacements for different diaphragm position differences within an interval of 2–20 mm and different cardiac phases between precontrast and contrast-enhanced CT images.

Diaphragm position difference	Cardiac phase (RR interval)	Median displacement (mm)	25th–75th–95th percentile range (mm)
3 mm	42%	0.78	0.36–1.82–3.09
8 mm	83%	1.94	0.87–4.39–7.40
14 mm	25%	3.53	1.54–8.05–13.72
19 mm	67%	4.70	2.02–10.74–18.31
20 mm	8.3%	5.04	2.14–11.51–19.59
19 mm	50%	4.75	2.03–10.89–18.60
16 mm	92%	4.01	1.73–9.08–15.41
12 mm	33%	3.03	1.32–6.90–11.76
8 mm	75%	1.93	0.87–4.36–7.35
4 mm	17%	0.95	0.44–2.18–3.67
2 mm	58%	0.36	0.18–0.85–1.45

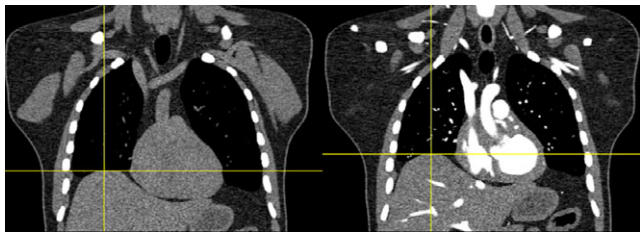

FIG. 2. Coronal slices of the XCAT phantom, after CT simulation (window width/window level (ww/wl): 500/80). The voxels are $1.0 \times 1.0 \times 1.0$ mm^3. (left) The precontrast scan at inspiration and (right) the contrast-enhanced scan at expiration. The difference in position of the dome of the diaphragm (yellow horizontal lines) is 20 mm. [Color figure can be viewed at wileyonlinelibrary.com]

in the lung parenchyma. Two nodules represented as spheres with a diameter of 5 mm and with attenuation values equivalent to liver tissue were added to the upper lung and close to the diaphragm. Movement of nodules between the different respiratory levels was simulated according to the movement of the center voxel in the nodules. The registration performance was tested at the same respiratory levels as previously, and the depiction of the nodules in the iodine maps for different enhancement levels and registration performance was evaluated by comparing the residual error of the location of nodule, with the size of the nodule. The influence of the presence of a wedge-shaped perfusion defect on the registration performance was evaluated by creating wedge-shaped perfusion in the upper and lower lungs. For this, the difference in pulmonary parenchyma enhancement between the wedge-shaped area representing a perfusion defect and the normally perfused pulmonary parenchyma was set to 10%. These perfusion defects were added in two locations; close to the diaphragm with a maximum diameter of 25 and 50 mm, and, in the upper lung with a radius of 25 mm. The influence on the registration of the algorithm of these three perfusion defects was studied for the 0, 8, and 20 mm diaphragm differences. Finally, the iodine map was generated from the registered and subtracted CT images to determine whether the perfusion defects were shown on the iodine map. To perform an initial evaluation of the impact of the resulting registration residual errors on the visibility of perfusion defects, their detectability after varying registration performance by the commercially available product was evaluated by one thoracic radiologist, blinded to their location.

2.C. Image simulation

The values of the voxels in the XCAT phantom represent linear attenuation coefficients at a specific user-defined x-ray energy. To generate realistic phantom images, phantom realizations for x-ray energies were generated from 10 to 100 keV, in 1 keV steps. Using the TASMICS software from Hernandez and Boone,[14] an x-ray spectrum with a tube potential of 100 kV and an added filter thickness of 17.4 mm aluminum was generated, resulting in a first half value layer (HVL) of 8.28 mm Al. This first HVL value was similar to measurements on the CT scanner in our radiology department (Aquilion ONE GENESIS edition, Canon Medical Systems, Otawara, Japan). The monoenergetic attenuation values of each voxel were used in a polychromatic simulation to generate realistic CT images. For this, the monoenergetic attenuation phantom data for each energy bin was forward projected and combined into polychromatic projections using the abovementioned normalized x-ray spectrum. Then, Poisson quantum noise was added to the projections and the logarithm was taken. Finally, a water precorrection was applied and an analytic filtered back projection reconstruction with a clinically relevant filter kernel was performed. The resulting polychromatic CT images were rescaled to Hounsfield Units (HU).

Initially, the level of quantum noise added was set to be equivalent to the noise variance found in ROIs in the liver in clinical CT scans. This resulted in the setting of the unattenuated photon count incident at each detector pixel being set at 8,000 photons. To investigate the influence of noise, the simulated radiation dose was then decreased by a factor of 10 and increased by a factor of 2.

2.D. Analysis

In addition to the phantom realizations, the XCAT program also outputs a map of 3D deformation vectors with the true voxel-by-voxel displacement between the two generated images. The displacement of each voxel was described in the X-, Y-, and Z-directions. This ground truth was compared to the results of each registration algorithm to obtain the voxel-by-voxel residual error both in each direction and in absolute terms. This analysis was performed only on the lung voxels, identified as such using the lung mask generated by the XCAT phantom. Although the individual-direction residual errors were obtained, in the results we presented only the absolute magnitude of the residual errors due to space and clarity constraints. These were shown using the median, 25th, 75th, and 95th percentile. For conciseness, histograms were shown only when results varied substantially with varying conditions. The calculations and analyses were performed in Matlab.

2.E. Image registration algorithms

Registration between the simulated precontrast and contrast-enhanced CT images was tested for three different algorithms. Algorithm A, CMRE NRR v1.0, is a commercially available registration- and motion-correction algorithm used in clinical practice (SURE Subtraction Lung, version 6.0, Canon Medical Systems, Otawara, Japan).[15] This algorithm is mainly developed to image pulmonary perfusion by visualizing the iodine enhancement in the lungs, which can show perfusion defects caused by pulmonary embolism. Algorithm B is CMRE NRR v2.0 of the same SURE Subtraction Lung algorithm, which is a prototype and hence not yet commercially available. It allows for tissue sliding at the interface between the lung tissue and the pleura during registration.[16] Algorithm C is the MEVIS DIS-CO algorithm (Fraunhofer-Gesellschaft, Lübeck, Germany),[17] which is designed for correction of large respiratory motion by integrating sparse keypoint correspondences into a dense continuous optimization framework, which placed first in the EMPIRE 10 challenge. This registration algorithm was developed for large motion differences, for example, from patients with chronic obstructive pulmonary diseases in whom accurate pulmonary registration leads to reliable quantification of local lung ventilation. Other applications are the estimation of tumor motion for radiotherapy treatment to save patient dose and the follow-up of lung nodules during lung screening.

All three algorithms perform their own lung segmentation prior to determination of the displacement field between precontrast and contrast-enhanced scans. No user-specified parameters were needed during the execution of the registration algorithm.

TABLE II. The median, interquartile range, and the 95th percentile of the residual errors of each registration algorithm for diaphragm position differences between 0 and 20 mm and different cardiac phases between precontrast and contrast-enhanced CT images.

Diaphragm position difference	Cardiac phase (RR interval)	Registration algorithm	Median residual error (mm)	25th–75th–95th percentile range (mm)
0 mm	0%	A	0	0–0–0
		B	0.29	0.19–0.44–1.06
		C	0.33	0.23–0.51–1.22
3 mm	42%	A	0.32	0.30–0.76–1.43
		B	0.26	0.17–0.39–1.34
		C	0.38	0.23–0.65–1.71
8 mm	83%	A	0.49	0.30–0.76–1.42
		B	0.49	0.32–0.75–1.67
		C	0.54	0.34–0.86–1.73
14 mm	25%	A	0.62	0.42–1.14–3.65
		B	0.70	0.37–1.05–2.97
		C	0.75	0.45–1.24–3.04
19 mm	67%	A	0.83	0.50–1.44–5.61
		B	0.81	0.49–1.37–3.50
		C	0.88	0.52–1.49–3.86
20 mm	8.3%	A	0.85	0.51–1.47–5.55
		B	0.82	0.50–1.40–3.59
		C	0.91	0.54–1.52–4.02
19 mm	50%	A	0.81	0.48–1.47–6.58
		B	0.79	0.48–1.39–3.96
		C	0.89	0.53–1.55–4.32
16 mm	92%	A	0.72	0.44–1.22–3.66
		B	0.66	0.41–1.10–2.56
		C	0.80	0.48–1.29–2.99
12 mm	33%	A	0.63	0.39–1.05–3.40
		B	0.58	0.36–0.97–2.76
		C	0.68	0.41–1.13–2.82
8 mm	75%	A	0.49	0.31–0.76–1.51
		B	0.48	0.31–0.75–1.71
		C	0.55	0.34–0.86–1.75
4 mm	17%	A	0.43	0.28–0.68–1.61
		B	0.28	0.19–0.42–1.13
		C	0.41	0.26–0.68–1.59
2 mm	58%	A	0.26	0.15–0.48–1.11
		B	0.25	0.17–0.38–1.21
		C	0.32	0.21–0.57–1.61

3. RESULTS

3.A. Influence of different diaphragm levels and cardiac phases

Table I shows the true displacements for the different diaphragm position differences and cardiac phases, during one respiratory cycle, between the precontrast and contrast-enhanced scan, while Table II shows the median, the interquartile ranges, and the 95th percentile of the residual errors for the three algorithms. While algorithm A did not introduce errors when images of the same respiratory phase were selected as input, algorithm B and C did estimate a deformation field, with a small median residual error. In general for all three algorithms, the residual error increased with larger diaphragm differences between the two scans. Figure 3 shows the histogram of the residual error per voxel for the 3, 8, and 20 mm diaphragm position difference resulting from algorithm A. The other algorithms show comparable graphs. The clinical average diaphragm difference (8 mm), showed a median residual error below 0.54 mm, implying that the median residual error was almost half of the voxel size of 1.0 mm. The distribution in residual errors for the 20 mm difference is shown in Fig. 4 for all three algorithms. Even for a large diaphragm position difference between the precontrast and contrast-enhanced CT scans, the median residual error was below 1 mm for all algorithms, again smaller than one voxel. The algorithms showed similar results, but overall algorithm B had the smallest residual error.

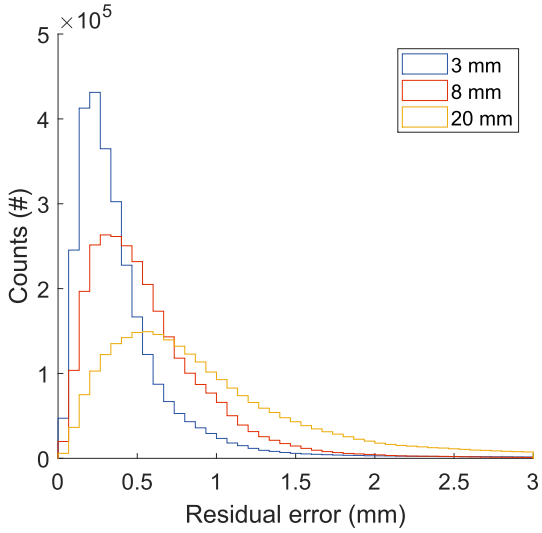

FIG. 3. Histograms of the residual error of three different diaphragm positions of algorithm A. [Color figure can be viewed at wileyonlinelibrary.com]

Figure 5 shows a 1 mm thick slice of the residual errors from the three algorithms at different positions in the lungs for the case with a diaphragm position difference of 20 mm. It can be seen that besides at the paracardiac regions, the region around the diaphragm does also present large errors. This section of the lungs undergoes the largest movement during breathing. The three tested algorithms show approximately the same residual error maps. Algorithm B shows the improvement of the registration algorithm compared to algorithm A in the paracardiac region, where algorithm C shows very small errors. The images look smooth because of the slow-varying difference in voxel motion across adjacent voxels.

3.B. Influence of whole body motion

Figure 6 shows the displacement in each direction when whole body motion was simulated. The results of the median residual error after registration with this whole

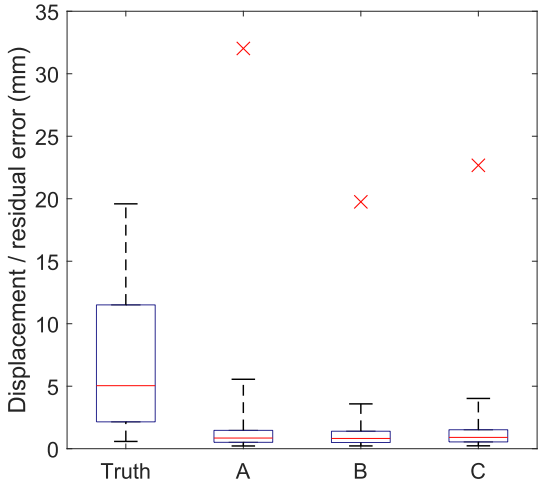

FIG. 4. The true displacement and the residual error of the displacement estimated by the three algorithms, for all voxels, tested for a diaphragm position difference of 20 mm between the two scans. The distribution shows the median (red line), the 25th–75th percentile (blue box), 5th–95th percentile (black whiskers), and the maximum outliers (red crosses). [Color figure can be viewed at wileyonlinelibrary.com]

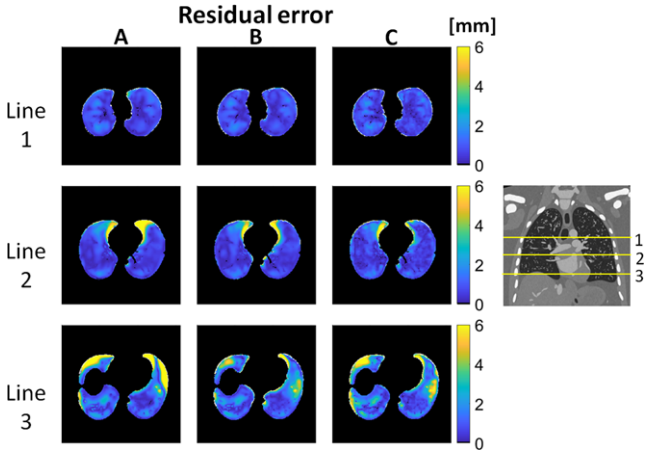

FIG. 5. The residual error of the three registration algorithms in axial slices at different levels of the lungs (approximate position depicted on the right) with a diaphragm difference of 20 mm between two scans. [Color figure can be viewed at wileyonlinelibrary.com]

body motion in each of the X, Y, and Z directions are shown in Table III, along with the corresponding results without this motion from Table II, for comparison purposes. In general, algorithm B seems to be affected by the presence of whole body motion, while the other two algorithms are not sensitive to this additional variation. This influence of each direction separately on the algorithm B is shown in Fig. 7.

3.C. Influence of iodine enhancement level

The mean enhancement, estimated over a circular region of interest on a single slice in the pulmonary truncus, was 275 HU with reduced iodine enhancement, 452 HU with standard enhancement, and 606 HU with high iodine enhancement. Figure 8 shows these three contrast-enhanced images with the same window width/ window level. Table IV lists the median residual errors when varying the iodine enhancement with minimal influence of iodine on the residual error of the three algorithms for the 8 and 20 mm diaphragm differences. With a reduction of iodinated contrast, algorithm B did not show any initial errors anymore when the same respiratory precontrast and contrast-enhanced level was used.

3.D. Influence of beam hardening artifacts

Figure 9 shows the titanium screws and rods in the spine (sagittal and axial plane) and the corresponding iodine map. The median residual errors of the registration algorithm in this situation are presented in Table V; the impact of beam hardening artifacts caused by the strong material on the median residual error is small, whereas the impact on the iodine map is high, as is visible in Fig. 9.

3.E. Influence of the presence of nodules in the lungs

The locations of the two added nodules are indicated in Fig. 10. Overall, the nodules have minimal influence on the registration algorithms, even reducing the resultant residual error in some cases (Table VI). Figure 11 shows the iodine maps of the nodules and the registration errors in the same area. Here, the residual error of location 1 near the diaphragm is larger than 5 mm for the 20 mm difference in diaphragm level, which is larger than the size of the nodule, while the residual error of location 2 (upper lobe) is around 1 mm. For the 8 mm difference in diaphragm level, both nodules are in areas with residual errors below 2 mm.

FIG. 6. Displacement in (left) X-, (middle) Y-, and (right) Z-direction in the same slice position (ww/wl: 1600/−600).

TABLE III. The median of the residual errors of each registration algorithm for diaphragm position differences 0, 8, and 20 mm between precontrast and contrast-enhanced CT images where a X, Y, or Z displacement was introduced at the precontrast scan.

Diaphragm position difference	Registration algorithm	Median residual errors (mm)			
		No whole body displacement	X displacement	Y displacement	Z displacement
0 mm	A	0	0.24	0.38	0.31
	B	0.29	0.44	0.33	0.36
	C	0.33	0.37	0.38	0.35
8 mm	A	0.49	0.51	0.49	0.49
	B	0.49	0.78	0.73	0.57
	C	0.54	0.55	0.54	0.55
20 mm	A	0.85	0.88	0.87	0.87
	B	0.82	1.27	1.34	0.97
	C	0.91	0.90	0.90	0.90

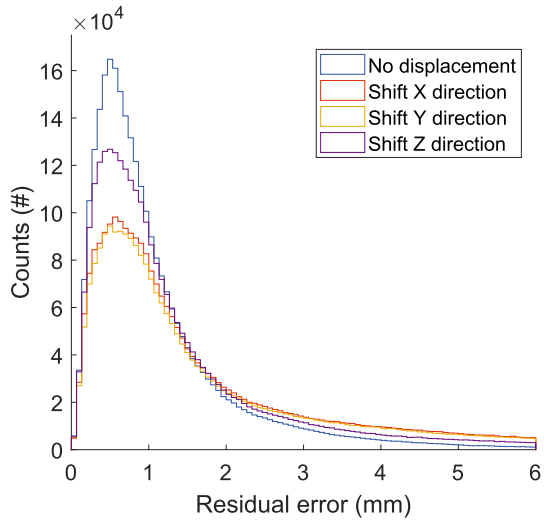

FIG. 7. Histograms of the residual errors of each voxel for the standard situation and for each of the three whole body movement changes at a diaphragm displacement of 20 mm between the precontrast and contrast-enhanced scan for algorithm B. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE IV. The median of the residual errors of each registration algorithm for diaphragm position differences 0, 8, and 20 mm between precontrast and contrast-enhanced CT images. The contrast-enhanced image had a reduced or an increased iodine enhancement compared to the standard situation.

Diaphragm position difference	Registration algorithm	Median residual errors (mm)		
		Standard concentration	Reduced concentration	Increased concentration
0 mm	A	0	0	0
	B	0.29	0	0.31
	C	0.33	0.33	0.33
8 mm	A	0.49	0.47	0.51
	B	0.49	0.47	0.51
	C	0.54	0.55	0.55
20 mm	A	0.85	0.84	0.87
	B	0.82	0.86	0.83
	C	0.91	0.90	0.90

3.F. Influence of the presence of perfusion defects

The influence of the perfusion defects on the registration algorithms can be seen in Table VII, and it shows a similar impact as the presence of nodules. There is a minimal difference in the median residual error of the registration algorithms compared to the images without perfusion defects. Figure 12 shows the results of the iodine map for all locations using algorithm A, with a 20 mm difference. It indicates that a perfusion defect with a diameter of 50 mm is visible in the lower lobe (location 1), an area with a high median residual error of 7.66 mm. Smaller perfusion defects (25 mm) in the same area (location 2) were not visible, while they were equivocally detectable when placed in the upper lobe (location 3), where the median residual error was only 0.85 mm.

FIG. 8. (Left) The reduced iodine contrast concentration, (middle) the standard concentration, and (right) the increased concentration (ww/wl: 750/90) .

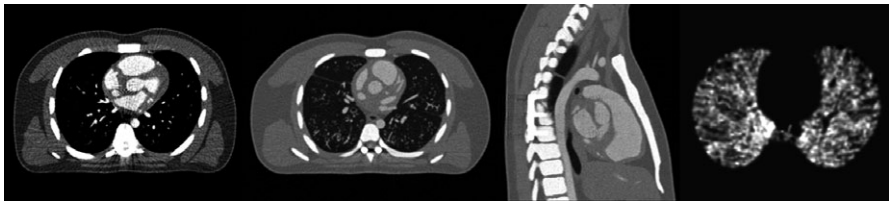

FIG. 9. From left to right: Titanium screws and rods with beam hardening artifacts on the axial image in ww/wl: 90/750, axial image on another slice, to show the dark line between the rods with ww/wl: 0/2000, and sagittal the pins T6 to T8, and the corresponding (right) iodine map.

3.G. Influence of noise

Figure 13 illustrates the three different noise levels that were obtained by changing the number of incident photons in the CT simulations. The median residual errors are shown in Table VIII. Algorithm B was insensitive to increased noise, while algorithm A showed changes within 0.04 mm, and for algorithm C, the impact on the error was larger but decreased with increased difference in respiratory level. There is also no clear reduction visible in residual error when the dose is increased (<0.07 mm), except for algorithm B that did not introduce any initial errors when the diaphragm levels were equal.

4. DISCUSSION

In a situation that simulates the average motion in a normal breathing patient (8 mm diaphragm difference), the current commercial subtraction software could compensate for respiratory motion within a median error of 0.49 mm, similar to the other state-of-the-art algorithms. This implies that motion correction in subtraction imaging is accurate enough to detect and characterize large area differences of iodine enhancement in the lungs of at least 50 mm and larger, such as for the detection of perfusion defects in patients with pulmonary embolism, without much negative influence of major registration errors caused by breathing. The success of the algorithms for perfusion defects between 25 and 50 mm depends on their location. Areas that have small registration errors will depict these defects, but otherwise they will not be visible. We did not investigate smaller perfusion defects (<25 mm), in detail, but most likely they will often not be shown correctly and therefore cannot be detected. As mentioned, subsegmental perfusion defects are in the centimeter range,[11] which explains why they are reported as usually visible in previous studies.[1,18–20] Furthermore, since the appropriateness of treating very small subsegmental embolisms is debatable, in cases where these perfusion defects are not depicted, this might not have a deleterious impact on the diagnosis and determination of prognosis of PE.[21,22] The perfusion maps may be less accurate, for structures in the paracardiac, above the diaphragm, and ventral subpleural regions, due to larger movement caused by breathing. This could be checked in clinical practice by generating minimal intensity projections of the iodine maps. In general, larger diaphragm differences between the precontrast and contrast-enhanced scans create larger errors. For large diaphragm position differences, the error could be of magnitudes equivalent to one or two voxels. Beam hardening artifacts have major impact on the iodine map, but not on the residual error of the registration algorithms. A severe increase in noise in the image affects the results of the registration algorithm, but the impact in registration errors for the iodine map is smaller than that on the overall precontrast and contrast-enhanced image quality, which makes the image inappropriate for diagnosis. All other simulations with differences in body movement, and iodine enhancement, did not have a major relevant impact on the accuracy of the registration algorithms for large area differences in iodine distribution. Registration for the detection of enhancement in pulmonary nodules smaller than

TABLE V. The median of the residual errors of each registration algorithm for diaphragm position differences 0 mm, 8 mm, and 20 mm between precontrast and contrast-enhanced CT images, without and with beam hardening artifacts caused by titanium screws and rods in the spine.

Diaphragm position difference	Registration algorithm	Median residual errors (mm)	
		No beam artifacts	Beam artifacts
0 mm	A	0	0
	B	0.29	0.28
	C	0.33	0.31
8 mm	A	0.49	0.50
	B	0.49	0.49
	C	0.54	0.56
20 mm	A	0.85	0.87
	B	0.82	0.90
	C	0.91	0.92

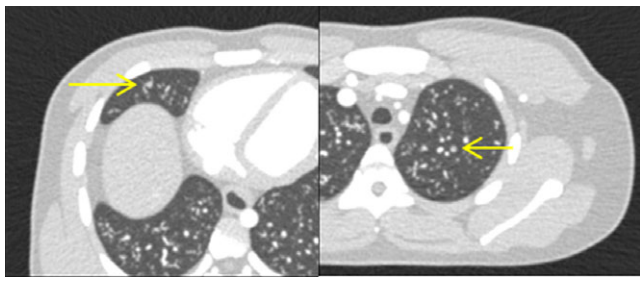

FIG. 10. The locations of the nodules (arrows) in the contrast-enhanced images on lung window (ww/wl: 1600/−600), (left) in the lower lobe, nearby the diaphragm — location 2, and (right) in the upperlobe — location 1. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE VI. The median of the residual errors of each registration algorithm for diaphragm position differences of 0 mm, 8 mm, and 20 mm between precontrast and contrast-enhanced CT images, without and with nodules.

Diaphragm position difference	Registration algorithm	Median residual errors (mm)		
		Without nodules	Nodule location 1	Nodule location 2
0 mm	A	0	0	0
	B	0.29	0.12	0.31
	C	0.33	0.33	0.32
8 mm	A	0.49	0.48	0.49
	B	0.49	0.49	0.49
	C	0.54	0.54	0.54
20 mm	A	0.85	0.87	0.86
	B	0.82	0.83	0.78
	C	0.91	0.90	0.90

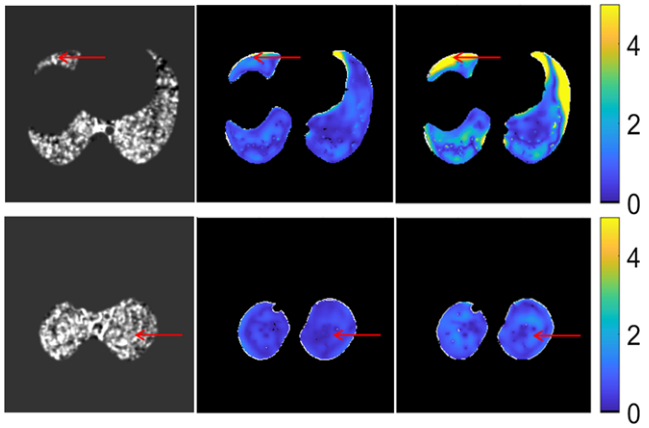

FIG. 11. (Left) The two iodine maps of algorithm A and the corresponding residual error map with (middle) 8 mm and (right) 20 mm difference in diaphragm level. The 5 mm sphere close to the diaphragm is located (red arrow) in an area with a residual error larger than the diameter of the nodule for the 20 mm, but not for the 8 mm. The nodule in the upper lobes lays in the residual error that shows smaller values than the nodule diameter. [Color figure can be viewed at wileyonlinelibrary.com]

5 mm is only reliable in the upper half of the lungs (avoiding the area around the heart) for a 20 mm difference between the diaphragm levels. At this diaphragm level difference, in the lower part of the lungs, the residual error is larger than the size of the nodule, making the detection of pulmonary nodule enhancement may not be reliable. However, for average clinical scans, with a difference in diaphragm level of 8 mm, the registration for the detection of enhancement of pulmonary nodules is reliable except for the paracardiac region and the regions closest to the diaphragm.

For evaluation of this type of motion correction algorithms, it is important to use a phantom with anatomical structures, representing vessels and airways, of similar spatial distribution as encountered in clinical practice. This is due to the fact that the algorithms use the structures inside the lungs as references for their registration. Hence, the use of the newest version of the XCAT phantom, which has been reported

TABLE VII. The median of the residual errors of each registration algorithm for diaphragm position differences of 0, 8, and 20 mm between precontrast and contrast-enhanced CT images, without and with perfusion defects in the contrast-enhanced image. Location 1: radius of 25 mm close to the diaphragm, location 2: radius of 12.5 mm close to the diaphragm, and location 3: radius of 50 mm in the apex.

Diaphragm position difference	Registration algorithm	Median residual errors (mm)			
		Without PD	PD location 1	PD location 2	PD location 3
0 mm	A	0	0	0	0
	B	0.29	0.27	0.32	0.33
	C	0.33	0.33	0.33	0.33
8 mm	A	0.49	0.49	0.50	0.49
	B	0.49	0.52	0.51	0.51
	C	0.54	0.54	0.54	0.55
20 mm	A	0.85	0.86	0.86	0.88
	B	0.82	0.83	0.79	0.83
	C	0.91	0.88	0.88	0.90

to be closer to clinical reality in terms of the airways and vessels than the previous version,[13] was important to maximize the relevance of this study.

These phantom simulations enable objective comparison of registration algorithms by providing a known ground truth. This study demonstrated that the newer Canon registration algorithm shows an, albeit small, improvement in performance compared to the current commercially available algorithm, especially in the paracardiac region. Algorithms A and B were tested in previous studies with different methods and smaller diaphragm motion,[15,16] with similar results. However, in those works, the authors evaluated the residual error in each direction individually instead of its absolute magnitude, and only reported the mean and standard deviation in estimating the movement of 100 landmarks rather than evaluating all voxels included in the lungs, as done in this work. The MEVIS DIS-CO algorithm is used for registration of chest CT images acquired longitudinally for nodule follow-up, but this study shows that it is also appropriate for use in subtraction CT for pulmonary perfusion evaluation, resulting in good registration of the two scans and a better performance in the paracardiac regions. Although the prototype algorithm shows better results when the diaphragm changes position due to respiratory motion, the presence of whole body movement had a negative impact on its performance, especially in the lateral and craniocaudal directions.

Our study has a few limitations. First, we investigated the accuracy of the registration algorithms with synthetic digital phantom images instead of real patient images. However, this gave us the major advantage that we have a known ground truth for each voxel, which is impossible with real patient data. Therefore, this is the only method possible to investigate, objectively, the performance of each registration algorithm.

Second, we only investigated motion correction for differences in respiratory phase that reflected the vast majority (98%) of clinical cases. Since residual error increases with

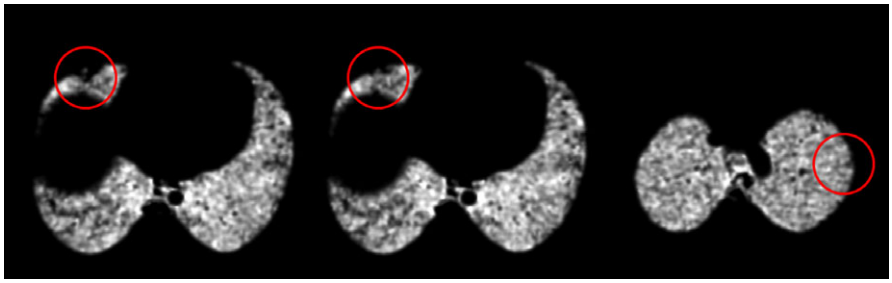

FIG. 12. Perfusion defects on the iodine map on (left) location 1, with a diameter of 50 mm, (middle) same location but with a perfusion defect of 25 mm that is invisible mainly due to other small perfusion defects, and (right) of location 3, with a diameter of 25 mm. All images have a difference in precontrast and contrast-enhanced scans of 20 mm and used algorithm A (ww/wl: 60/30). [Color figure can be viewed at wileyonlinelibrary.com]

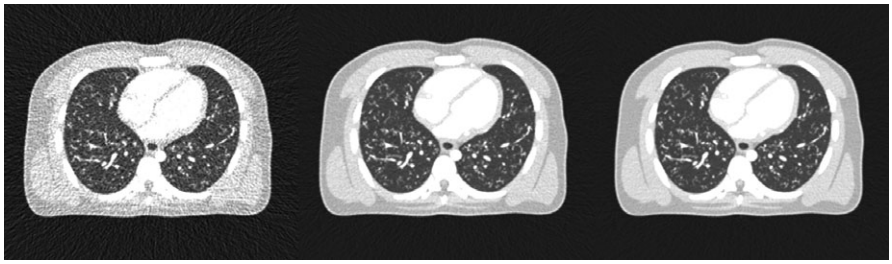

FIG. 13. (left) 0.1× of the original dose, (middle) original dose, (right) 2× the original dose. All images are 1 mm thick and have a ww/wl of −600/1600.

TABLE VIII. The median of the residual errors of each registration algorithm for diaphragm position differences 0 mm, 8 mm, and 20 mm between precontrast and contrast-enhanced CT images, with different levels of noise.

		Median residual errors (mm)		
Diaphragm position difference	Registration algorithm	Standard noise level	Decreased noise (dose ×2)	Increased noise (dose ×0.1)
0 mm	A	0	0	0
	B	0.29	0	0.29
	C	0.33	0.26	0.57
8 mm	A	0.49	0.48	0.52
	B	0.49	0.46	0.49
	C	0.54	0.51	0.69
20 mm	A	0.85	0.86	0.89
	B	0.82	0.83	0.82
	C	0.91	0.88	1.04

larger differences in diaphragm position, it is likely that extreme motion will lead to larger registration inaccuracies.

Third, the phantom simulates respiratory and cardiac motion. It was not possible to add vascular pulsation, but we did not consider this an issue, since these pulsations are small compared to the other two sources of motion, and would probably not influence the registration performance much.

Finally, a male digital phantom was used that was based on scans of healthy volunteers. The motion for patients suffering disease might be different. Although this phantom is an excellent tool for this purpose, this "healthy" phantom does not reflect the situation of patients with significant chest abnormalities, such as atelectasis, pleural fluid, consolidations, and chest deformities. It is unknown what the effect of these pathologies would be on the performance of the registration algorithms.

The median residual error of less than a single voxel suggests that with further improvement around the heart, and for larger differences in diaphragm levels that are above clinical average, they could, in the future, also be used for temporal subtraction in multiphase imaging for pulmonary nodule characterization and follow-up, and for assessment of pulmonary inflammation. For pulmonary nodules of 5 mm or larger, at clinical relevant diaphragm levels between the precontrast and contrast-enhanced scan, the residual error is below 2 mm. Therefore, for pulmonary nodules 5 mm diameter or larger located away from the paracardiac and the diaphragm regions, the use of registration for subtraction CT is reliable.

For readers who would be interested in testing their own registration algorithms on these images and compare their performance to the results presented here, these have been made publicly available in a new challenge that can be found at https://lumic.grand-challenge.org/.

5. CONCLUSION

Registration algorithms for temporal subtraction CT imaging can accurately compensate for respiratory motion between scans. Therefore, at least from the point of view of differences in breathing, cardiac phase, amount of iodine contrast, beam hardening artifacts, and whole patient movement, iodine maps for evaluation of vascular obliteration sequelae can be generated reliably in most patients, with a median residual error that is smaller than one voxel. These conditions

do not cause large errors in registration, but they can influence the resulting iodine map, especially the beam hardening artifacts. Perfusion defects such as those caused by pulmonary embolism that is larger than 50 mm will be depicted in the iodine maps, even if median registration errors of 7.7 mm are present in the specific area.

ACKNOWLEDGMENT

We thank W.P. Segars (Duke University, North-Carolina, USA) for the use and support with the XCAT phantom. Parts of the reconstruction software were provided by RayConStruct® GmbH, Nürnberg, Germany.

CONFLICTS OF INTEREST

Funding for this research has been provided by Canon Medical Systems Corporation, manufacturer of CT systems, and developers of two of the algorithms evaluated in this study. The study data and results were generated and controlled at all times by the research personnel at Radboud University Medical Centre, with no influence from Canon.

[a)]Author to whom correspondence should be addressed. Electronic mail: Ioannis.Sechopoulos@radboudumc.nl.

REFERENCES

1. Pontana F, Faivre JB, Remy-Jardin M, et al. Lung perfusion with dual-energy multidetector-row CT (MDCT): feasibility for the evaluation of acute pulmonary embolism in 117 consecutive patients. *Acad Radiol*. 2008;15:1494–1504.
2. Thieme SF, Johnson TR, Lee C, et al. Dual-energy CT for the assessment of contrast material distribution in the pulmonary parenchyma. *AJR Am J Roentgenol*. 2009;193:144–149.
3. Ohno Y, Nishio M, Koyama H, et al. Solitary pulmonary nodules: comparison of dynamic first-pass contrast-enhanced perfusion area-detector CT, dynamic first-pass contrast-enhanced MR imaging, and FDG PET/CT. *Radiology*. 2015;274:563–575.
4. Moon JW, Bae JP, Lee HY, et al. Perfusion- and pattern-based quantitative CT indexes using contrast-enhanced dual-energy computed tomography in diffuse interstitial lung disease: relationships with physiologic impairment and prediction of prognosis. *Eur Radiol*. 2016;26:1368–1377.
5. Grob D, Oostveen LJ, Prokop M, Schaefer-Prokop CM, Sechopoulos I, Brink M. Imaging of pulmonary perfusion using subtraction CT angiography is feasible in clinical practice. *Eur Radiol*. 2019;29:1408–1414.
6. Grob D, Smit EJ, Oostveen LJ, et al. Image quality of iodine maps for pulmonary embolism: a comparison of subtraction CT and dual-energy CT. *AJR Am J Roentgenol*. 2019;212:1–7.
7. Wildberger JE, Klotz E, Ditt H, Spuntrup E, Mahnken AH, Gunther RW. Multislice computed tomography perfusion imaging for visualization of acute pulmonary embolism: animal experience. *Eur Radiol*. 2005;15:1378–1386.
8. Tustison NJ, Cook TS, Song G, Gee JC. Pulmonary kinematics from image data: a review. *Acad Radiol*. 2011;18:402–417.
9. Murphy K, van Ginneken B, Reinhardt JM, et al. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. *IEEE Trans Med Imaging*. 2011;30:1901–1920.
10. Kirby N, Chuang C, Ueda U, Pouliot J. The need for application-based adaptation of deformable image registration. *Med Phys*. 2013;40:011702.
11. Screaton NJ, Coxson HO, Kalloger SE, et al. Detection of lung perfusion abnormalities using computed tomography in a porcine model of pulmonary embolism. *J Thorac Imaging*. 2003;18:14–20.
12. Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BM. 4D XCAT phantom for multimodality imaging research. *Med Phys*. 2010;37:4902–4915.
13. Abadi E, Segars WP, Sturgeon GM, Roos JE, Ravin CE, Samei E. Modeling Lung Architecture in the XCAT series of phantoms: physiologically based airways, arteries and veins. *IEEE Trans Med Imaging*. 2018;37:693–702.
14. Hernandez AM, Boone JM. Tungsten anode spectral model using interpolating cubic splines: unfiltered x-ray spectra from 20 kV to 640 kV. *Med Phys*. 2014;41:042101.
15. Goatman K, Plakas C, Schuijf JD, Beveridge E, Prokop M. Computed tomography lung iodine contrast mapping by image registration and subtraction. Proceedings of SPIE. 2014;9034.
16. Mohr B, Brink M, Oostveen LJ, Schuijf JD, Prokop M. Lung iodine mapping by subtraction with image registration allowing for tissue sliding. Paper presented at: SPIE Medical imaging 2016.
17. Rühaak J, Polzin T, Heldmann S, et al. Estimation of large motion in lung CT by integrating regularized keypoint correspondences into dense deformable registration. *IEEE Trans Med Imaging*. 2017;36:1746–1757.
18. Tamura M, Yamada Y, Kawakami T, et al. Diagnostic accuracy of lung subtraction iodine mapping CT for the evaluation of pulmonary perfusion in patients with chronic thromboembolic pulmonary hypertension: correlation with perfusion SPECT/CT. *Int J Cardiol*. 2017;243:538–543.
19. Weidman EK, Plodkowski AJ, Halpenny DF, et al. Dual-energy CT angiography for detection of pulmonary emboli: incremental benefit of iodine maps. *Radiology*. 2018;289:546–553.
20. Okada M, Kunihiro Y, Nakashima Y, et al. Added value of lung perfused blood volume images using dual-energy CT for assessment of acute pulmonary embolism. *Eur J Radiol*. 2015;84:172–177.
21. Wiener RS, Schwartz LM, Woloshin S. When a test is too good: how CT pulmonary angiograms find pulmonary emboli that do not need to be found. *BMJ*. 2013;347:f3368.
22. Peiman S, Abbasi M, Allameh SF, Asadi Gharabaghi M, Abtahi H, Safavi E. Subsegmental pulmonary embolism: a narrative review. *Thromb Res*. 2016;138:55–60.