BJR

Received: 15 October 2014 | Revised: 25 November 2014 | Accepted: 1 December 2014

doi: 10.1259/bjr.20140683

Cite this article as:
Skornitzke S, Fritz F, Klauss M, Pahn G, Hansen J, Hirsch J, et al. Qualitative and quantitative evaluation of rigid and deformable motion correction algorithms using dual-energy CT images in view of application to CT perfusion measurements in abdominal organs affected by breathing motion. Br J Radiol 2015;88:20140683.

FULL PAPER

Qualitative and quantitative evaluation of rigid and deformable motion correction algorithms using dual-energy CT images in view of application to CT perfusion measurements in abdominal organs affected by breathing motion

[1]S SKORNITZKE, MSc, [1]F FRITZ, MD, [1]M KLAUSS, MD, [1]G PAHN, MSc, [1]J HANSEN, MSc, [2]J HIRSCH, PhD, [1]L GRENACHER, MD, [1]H-U KAUCZOR, MD and [1]W STILLER, PhD

[1]Clinic of Diagnostic and Interventional Radiology, University Hospital Heidelberg, Heidelberg, Germany
[2]Department of Medical Biometry and Epidemiology (IMBE), University of Witten/Herdecke, Witten, Germany

Address correspondence to: Dr Wolfram Stiller
E-mail: *wolfram.stiller@med.uni-heidelberg.de*

Objective: To compare six different scenarios for correcting for breathing motion in abdominal dual-energy CT (DECT) perfusion measurements.

Methods: Rigid [RR_{Comm}(80 kVp)] and non-rigid [NR_{Comm}(80 kVp)] registration of commercially available CT perfusion software, custom non-rigid registration [NR_{Custom}(80 kVp], demons algorithm) and a control group [CG(80 kVp)] without motion correction were evaluated using 80 kVp images. Additionally, NR_{Custom} was applied to dual-energy (DE)-blended [NR_{Custom}(DE)] and virtual non-contrast [NR_{Custom}(VNC)] images, yielding six evaluated scenarios. After motion correction, perfusion maps were calculated using a combined maximum slope/Patlak model. For qualitative evaluation, three blinded radiologists independently rated motion correction quality and resulting perfusion maps on a four-point scale (4 = best, 1 = worst). For quantitative evaluation, relative changes in metric values, R^2 and residuals of perfusion model fits were calculated.

Results: For motion-corrected images, mean ratings differed significantly [NR_{Custom}(80 kVp) and NR_{Custom}(DE), 3.3; NR_{Comm}(80 kVp), 3.1; NR_{Custom}(VNC), 2.9; RR_{Comm}(80 kVp), 2.7; CG(80 kVp), 2.7; all $p < 0.05$], except when comparing NR_{Custom}(80 kVp) with NR_{Custom}(DE) and RR_{Comm}(80 kVp) with CG(80 kVp). NR_{Custom}(80 kVp) and NR_{Custom}(DE) achieved the highest reduction in metric values [NR_{Custom}(80 kVp), 48.5%; NR_{Custom}(DE), 45.6%; NR_{Comm}(80 kVp), 29.2%; NR_{Custom}(VNC), 22.8%; RR_{Comm}(80 kVp), 0.6%; CG(80 kVp), 0%]. Regarding perfusion maps, NR_{Custom}(80 kVp) and NR_{Custom}(DE) were rated highest [NR_{Custom}(80 kVp), 3.1; NR_{Custom}(DE), 3.0; NR_{Comm}(80 kVp), 2.8; NR_{Custom}(VNC), 2.6; CG(80 kVp), 2.5; RR_{Comm}(80 kVp), 2.4] and had significantly higher R^2 and lower residuals. Correlation between qualitative and quantitative evaluation was low to moderate.

Conclusion: Non-rigid motion correction improves spatial alignment of the target region and fit of CT perfusion models. Using DE-blended and DE-VNC images for deformable registration offers no significant improvement.

Advances in knowledge: Non-rigid algorithms improve the quality of abdominal CT perfusion measurements but do not benefit from DECT post processing.

CT perfusion measurements allow for the assessment of blood flow in the human body by injecting the patient with an iodinated contrast agent (CA) and tracking the changes in CT values in an organ, tissue or region of interest (ROI) over time. Mathematical perfusion models relate the measured time–enhancement curves to the physiological parameters governing blood flow, such as flow rate or local blood volume. This information can be used for assessment of diseases that influence local blood flow, for example, stroke or tumours.[1–3] As a result, CT perfusion is of particular interest for the detection of isodense tumours that show a low contrast to the surrounding tissue in the arterial and venous enhancement phase of conventional CT images, *e.g.* pancreatic tumours.[4]

While breathing motion of a patient during image acquisition has little influence on stroke assessment in the brain, CT perfusion acquisitions of the lung or abdominal organs

are usually performed during breath-hold or with breathing instructions (shallow breathing) to reduce the severity of breathing motion, depending on the type of perfusion model.[3] Abdominal perfusion protocols can take up to 2 min, which is much longer than the breath-hold duration to be expected from patients.[5] Therefore, abdominal organs will always be affected by breathing-induced motion during dynamic CT acquisition, diminishing the quality of the perfusion measurements.[3]

Motion correction can be performed by registering the CT images to a common reference image, eliminating changes in position induced by breathing motion (Figure 1). In order to achieve this, an optimization problem is solved to obtain a spatial transformation between the images by minimizing a measure for the differences in alignment. These spatial transformations are mainly classified as either rigid transformations (*e.g.* translation, rotation) or non-rigid transformations (*e.g.* deformation fields, free-form deformations).[6]

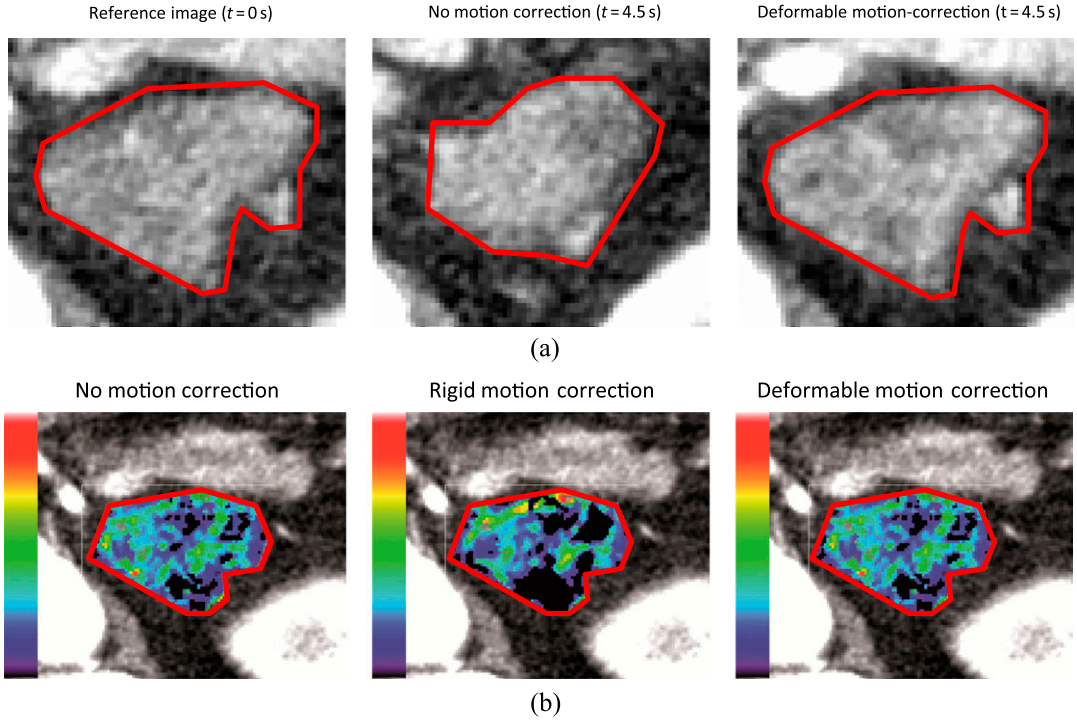
The use of a motion correction algorithm can be regarded as standard in current pancreatic perfusion studies.[4,5,7] While many studies have been performed to evaluate the performance of different registration algorithms for specific problems, the effect of different registration algorithms for correction of breathing motion in abdominal CT perfusion measurements has not yet been researched in detail (Figure 1).[6,8] Previous studies suggest that the use of non-rigid registration algorithms can improve the spatial alignment of lung and liver tumours during CT perfusion studies.[9,10]

The performance of motion-correction algorithms might be improved by the use of dual-energy CT (DECT) image data. DECT uses two acquisitions at different energy levels (*e.g.* 80 and 140 kVp) to generate additional information from the differences in absorption behaviour.[11] This allows for the calculation of virtual non-contrast (VNC) images that subtract the enhancement caused by the CA, ideally resulting in constant CT values within the ROI over the course of the whole perfusion measurement.[12] Furthermore, linearly blended dual-energy (DE) images with a 120-kVp-equivalent weighting factor of 0.5 can provide a higher contrast-to-noise ratio than can the 80-kVp images commonly used in CT perfusion measurements and might reduce the influence of image noise on motion correction.[13]

The purpose of this study was the assessment of the effectiveness of different algorithms for the correction of breathing motion in abdominal CT perfusion measurements in a clinical setting. Previous studies relying on a qualitative evaluation already reported a benefit in using non-rigid algorithms for the correction of breathing motion in liver tumours.[9] While quantitative methods, such as percentage overlap or distance of the centre of mass, have been suggested for the evaluation of motion-correction algorithms, these methods are limited by their need for segmentation of the motion-corrected images.[10]

In addition to the qualitative evaluation by experienced radiologists, this study provides a quantitative evaluation based on the optimization process of the motion correction and the fit of the perfusion models. Furthermore, the value of additional DECT post processing for the use in conjunction with deformable motion-correction algorithms was assessed. To our knowledge, this is the first study to provide a qualitative and

Figure 1. (a) Example of the effect of breathing motion and motion correction on the spatial alignment of the region of interest (ROI) in abdominal organs, *e.g.* a recurrent pancreatic tumour. Note the increasing iodine enhancement within the ROI already visible at $t = 4.5$ s. (b) Effect of breathing motion and motion correction on the calculation of perfusion maps.



quantitative evaluation of the effectiveness of different algorithms for the correction of breathing motion for CT perfusion measurements in recurring pancreatic tumours and the first study to include DECT image data for the purpose of motion correction.

METHODS AND MATERIALS

Patient collective

The study was approved by the local ethics committee and conducted in accordance with the ethical standards of the World Medical Association (Declaration of Helsinki). All patients gave written informed consent for study participation.

A total of 12 patients with suspected recurrent pancreatic carcinoma were included. The exclusion criteria for the CT perfusion acquisition were the following: manifested hyperthyreosis, decreased kidney function, inability to reproduce the breathing technique, known hypersensitivity to iodine CA and denial of consent.

Image acquisition

Image data sets were acquired with a dual-source DECT scanner (Somatom® Definition Flash; Siemens Healthcare, Forchheim, Germany). For tumour localization, a contrast-enhanced three-phase abdominal CT acquisition was performed (Table 1). After a washout period of 15 min, a CA bolus of 80 ml was applied at a flow rate of 5 ml s^{-1}, followed by a saline solution (NaCl) chaser bolus of 40 ml. The dynamic acquisition was initiated with a delay of 4 s after the start of the injection and consisted of 34 acquisitions over the course of 51 s at tube voltages of 80 and 140 kVp with tin filtration (Table 1). Using the appropriate conversion factor for abdominal CT examinations (0.015 mSv mGy^{-1} cm^{-1}), calculated average effective dose of the examinations was 16.8 mSv, subdivided in 7.8 mSv for the perfusion sequence and 9.1 mSv for the three-phase protocol.[14]

Motion correction

Three different motion correction algorithms, two from commercially available CT perfusion software (see below) and one developed in-house, were used to register the data sets of each patient to one reference image selected by an experienced radiologist. The algorithm developed in-house was implemented using the Insight Segmentation and Registration Toolkit (ITK) library and the ROOT library (CERN, Meyrin, Switzerland).[15,16] All algorithms used 80-kVp images for registration within a polygonal ROI set by an experienced radiologist. Additionally, the algorithm developed in-house was also applied to linearly blended DE images and VNC images, yielding a total of six evaluated scenarios, including the control group. Registration was limited to a rectangular image section fully containing the ROI to reduce computation times and to reduce the influence of other tissues and organs on the registration:

(1) Commercial rigid motion correction algorithm [RR_{Comm}(80 kVp); BodyPCT, Siemens Healthcare]. The BodyPCT software uses a two-step three-dimensional translation: the first step performs a correction in the axial plane, while the second step performs a correction along the craniocaudal axis. The displacement is limited to no more than 1 cm on each axis, and a mask is used to perform the registration solely on the basis of voxels with a CT value of below −19 HU. As technical limitations of the software prevent export of motion-corrected image data, the algorithm was recreated in our in-house-developed software according to the specifications given by the manufacturer. A visual comparison showed that the spatial alignment of images registered with the in-house reimplementation is comparable with that of images registered using the manufacturer workstation.

(2) Commercial non-rigid motion correction algorithm [NR_{Comm}(80 kVp); syngo®.via Body Perfusion, Siemens Healthcare]. The syngo.via Body Perfusion motion correction uses a smoothed deformation field. The motion correction was performed on the syngo.via workstation without definition of a ROI, as the deformation field is applied to the whole image volume. The motion-corrected data sets were exported for evaluation.

(3) Deformable motion correction algorithm [NR_{Custom}(80 kVp)]. A non-rigid registration using a smoothed deformation field was developed based on the ITK implementation of the Demons deformable algorithm.[17] The ITK software library implements a voxel-based version of the Demons deformable algorithm to calculate a deformation field based on the differences in image intensities, using a finite difference solver. By default, a histogram normalization filter is used prior to registration to reduce the differences in image intensities. As an addition to the ITK implementation, the deformable registration was constrained to a neighbourhood of the ROI and combined with a translation transformation for the initialization of the deformation field to reduce the computation time of

Table 1. CT acquisition parameters

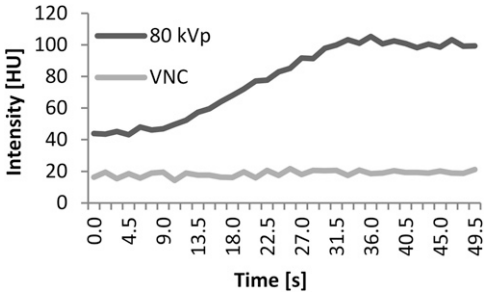
Phase	Tube voltage (kVp)	Effective tube-current time product (mAs)	Collimation (mm)	Reconstructed slice thickness (mm)	Delay (s)	Rotation time/cycle time (s)	Number of acquisitions
Native	120	200	2 × 64 × 0.6	3.0/1.5	4	0.5	Helical
Arterial	120	200	2 × 64 × 0.6	3.0/1.5	15	0.5	Helical
Venous	120	200	2 × 64 × 0.6	3.0/1.5	40	0.5	Helical
Perfusion	Sn140/80	50/270	32 × 0.6	3 × 5.0	4	0.5/1.5	34 dynamic

the algorithm. Additionally, a stopping criterion was implemented to stop the registration if the magnitude of the change of the average direction of the deformation field falls below a predefined threshold (see below).

(4) Deformable motion correction algorithm using DE-blended images [NR_{Custom}(DE)]. Linearly blended 120 kVp-equivalent DE images using a weighting factor of 0.5 were reconstructed on the CT scanner and registered using the deformable motion correction algorithm described above (see Point 3).
(5) Deformable motion correction algorithm using VNC images [NR_{Custom}(VNC)]. VNC images were calculated with software developed in-house and registered with the deformable motion-correction algorithm described above (see Point 3). For soft tissue and water, the necessary reference values for calculation of VNC images were taken from commercially available DE post-processing software (syngo. via Liver VNC; Siemens Healthcare) and from our own measurements for the iodine enhancement. An example for comparing mean CT values within the ROI for the VNC images to those of 80-kVp images over the course of the dynamic acquisition is given in Figure 2.
(6) Control group [CG(80 kVp)]. The original data set without any motion correction was used as a control group.

Where available, all algorithms used the default parameters provided by their respective vendors. The in-house registration software uses a mean squares metric and a regular step gradient descent optimizer as implemented in the ITK library with the exception of the reimplementation of the commercial rigid algorithm, where an optimizer supporting boundary conditions (limited memory Broyden Fletcher Goldfarb Shannon minimization with simple bounds) was used to enforce a movement no greater than 1 cm on each axis.[18] The stopping criterion for the deformable registration was set to a change in the average direction of the displacement field during an iteration of $<10^{-11}$ m, with an additional fixed maximum of no more than 10,000 iterations, which was determined experimentally to provide an adequate compromise between computation times and motion correction quality.

Figure 2. Comparison of mean CT values within a region of interest (ROI) over the course of the dynamic acquisition for the 80 kVp and virtual non-contrast (VNC) images. Note that, as expected, the mean CT value within the ROI does not change for the VNC images despite contrast agent injection.



Interpolation

While the segmentation was performed on the reconstructed data sets with a slice thickness of 5 mm, the motion correction was performed using 28 images with a slice thickness of 0.6 mm to take full advantage of the maximum resolution in *z* direction. As an exception, the reimplementation of the commercial rigid algorithm uses a slice thickness of 5 mm because of technical limitations imposed by the workstation.

After motion correction, slices with a thickness of 5 mm were linearly interpolated at the original positions from the motion-corrected data sets. Slices containing values from outside the acquisition volume inside the ROI were excluded from interpolation.

Calculation of perfusion maps

Colour-coded perfusion maps of the ROI were calculated from the motion-corrected image data using in-house developed software. All perfusion maps were calculated from 80-kVp images. Where the registration used linearly DE-blended or VNC images, the transformation obtained from the registration was applied to the 80-kVp images. The custom perfusion software implements a combined maximum slope and Patlak model for calculating colour-coded parameter maps for blood flow, permeability and blood volume, respectively.[19,20]

Motion correction and perfusion map quality evaluation

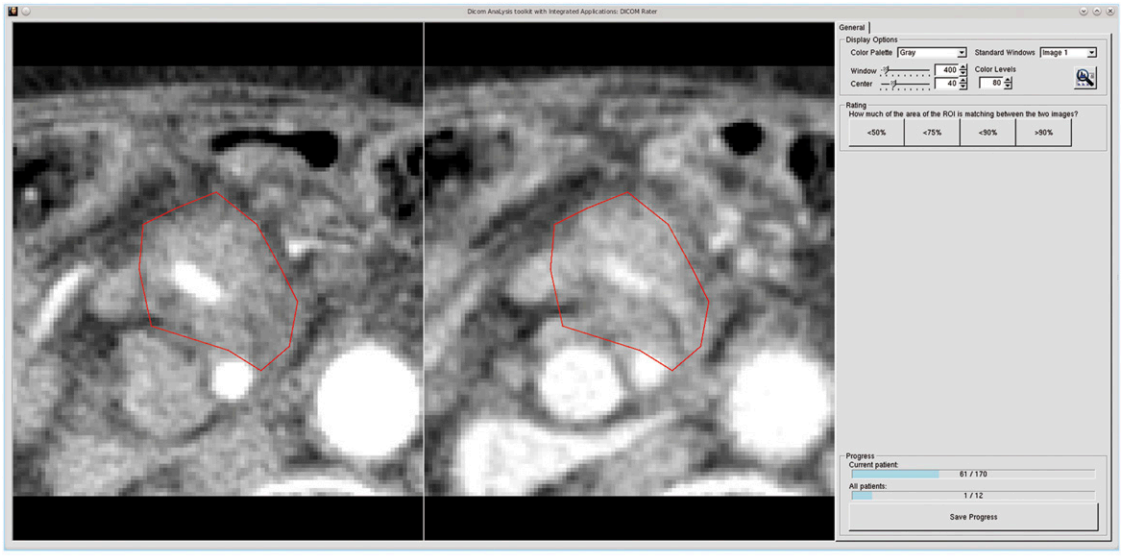
For qualitative evaluation, three blinded, experienced radiologists independently rated the quality of motion correction and resulting perfusion maps. A four-point scale (4 = best, 1 = worst) was agreed upon by the radiologists to properly reflect the impact of patient breathing motion on the image data. The scores were based on the area of the ROI unaffected by motion (unaffected area >90%; <90%; <75%; <50%). The parameter maps for the three perfusion parameters were each rated separately.

The rating was performed using the rating software developed in-house, providing anonymization and randomization. For the motion-corrected images, the rating software displays the reference image and the image that is to be rated side-by-side with an overlay of manual segmentation (Figure 3). For the perfusion maps, the rating software displays the perfusion maps side-by-side with the abdominal CT data sets used for tumour localization (native, arterial and venous phase; Table 1).

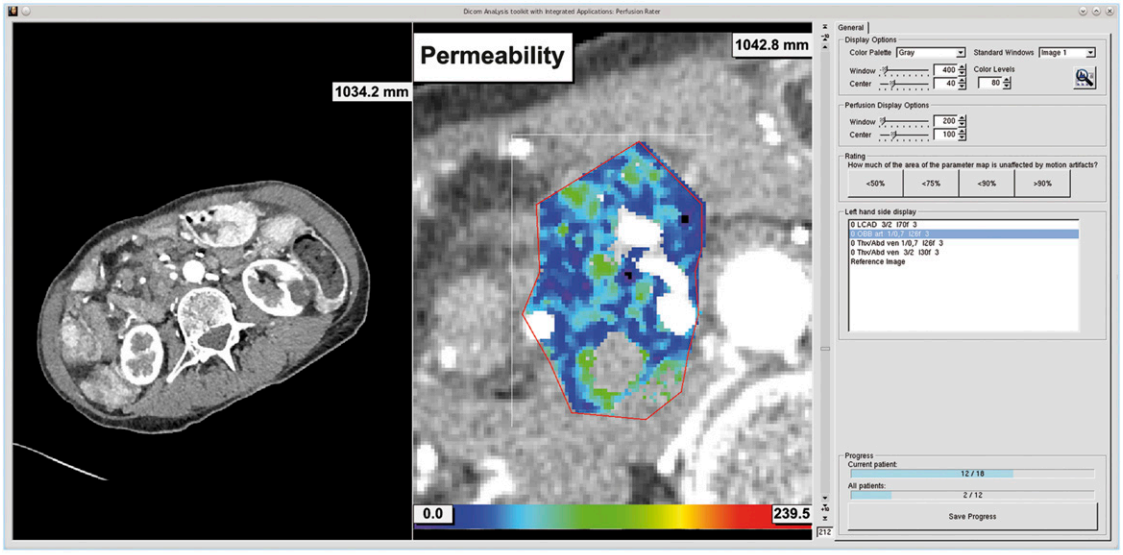
For the quantitative evaluation, the relative differences in the values of the mean squares metric before and after motion correction were evaluated, and the R^2 values and residuals of both the fit of the maximum slope and Patlak model were calculated for each voxel of the ROI.

For comparison of the qualitative and quantitative evaluation, the ratings of the motion-corrected images were first averaged between the three raters and the correlation between the absolute value of the mean squares metric after motion correction and the average rating of the image were calculated. Furthermore, the ratings of the respective perfusion maps were averaged

Figure 3. Screenshots of the rating software developed in-house: (a) rating software for the motion-corrected images, displaying the image to rate (right) side-by-side with the reference image (left). (b) Rating software for the colour-coded perfusion maps, displaying the perfusion maps (right) side-by-side with the reference image and abdominal CT acquisitions (left image; and selection on the right). The rating software provides randomization and anonymization, and allows for manual windowing as well as the selection of predefined windows.



(a)



(b)

between the three raters, and the correlation between the average rating and the average R^2 value and the residuals of the respective perfusion maps was calculated.

Statistics

A general linear model (GLM) was used to test for intergroup differences between the different algorithms. The GLM was applied to the ratings of the motion-corrected images and perfusion maps as well as the R^2 values and residuals of the maximum slope and Patlak model. The GLM modelled the three effects algorithm, patient and algorithm × patient, where patient and algorithm × patient are random effects [*i.e.* a linear model where the response variable (*e.g.* the rating) depends on the motion correction algorithm, the patient and all possible combinations of algorithms and patients].

The use of a GLM reduces the influence of the differences between patients, that is, the interindividual differences, as the severity of the breathing motion during the acquisition is individual. This correction is of even greater importance for the voxel-based evaluation of R^2 values and residuals of the fit of the maximum slope and Patlak models, as the size of ROI differs between patients as well.

Because of the large number of tests necessary for the pair-wise comparisons (comparison of 6 scenarios results in 15 tests), the

Holm–Bonferroni correction for multiple testing was used with a p-value of 0.05.

The partial Spearman correlation was used for measuring the correlation of the qualitative and quantitative evaluation, while correcting for the influence of the differences between patients.

RESULTS

The execution times of the commercial rigid and non-rigid algorithms were in the order of about 1 min (on the syngo.via server), while the registration with the deformable algorithms took between 20 min and 4 h (on an Intel® i7-X990 3.47-GHz central processing unit; Intel Corporation, Santa Clara, CA), depending on the severity of the breathing motion and the size of the ROI.

The GLM reported a significant effect ($p < 0.0001$) for the three effects algorithm, patient and algorithm × patient in all cases, indicating that these effects were correctly included in the model.

Motion-corrected images

On average, the images registered with deformable registration using 80-kVp or DE-blended images received the highest scores, followed by commercial non-rigid registration. The deformable registration using VNC images received worse scores than did the commercial non-rigid registration. The commercial rigid registration was on par with the control group (Figure 4a and Table 2). An example for the effect of the deformable motion correction can be seen in Figure 1a.

The ratings of the motion-corrected images showed statistically significant intergroup differences, with the exception of comparing deformable registration using 80-kVp and DE-blended images and comparing the commercial rigid registration to the control group (Table 3 part a).

The relative reduction in metric values owing to registration was the highest for deformable registration using 80-kVp or DE-blended images, followed by the commercial non-rigid registration and the deformable registration using VNC images. The commercial rigid registration achieved on average only a small relative reduction of the metric values (Figure 5e and Table 2).

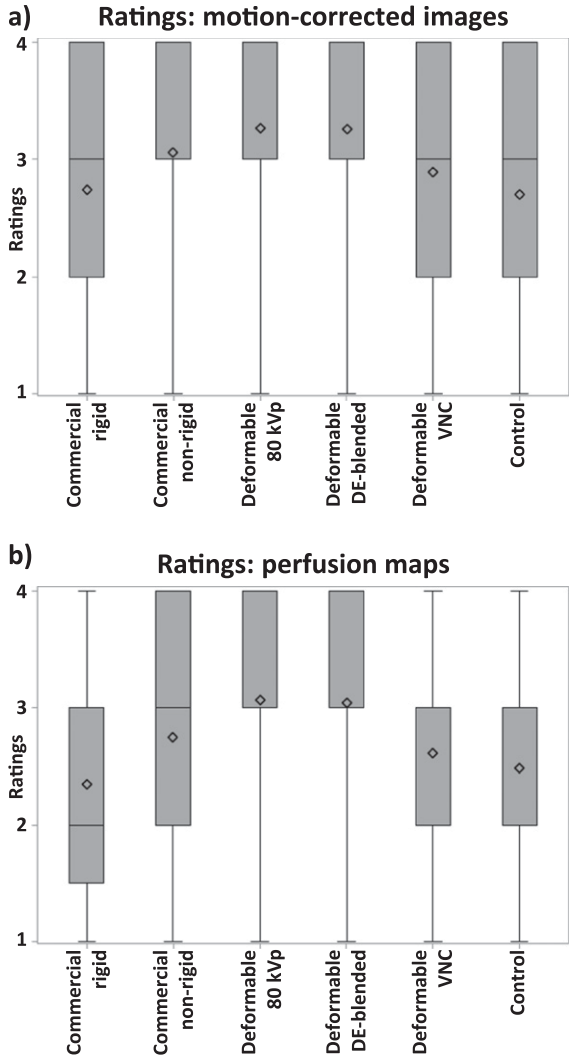
Perfusion maps

On average, the perfusion maps created after deformable 80-kVp or DE-blended registration received the highest scores, followed by the commercial non-rigid registration. The deformable registration using VNC images received worse scores than did the commercial non-rigid registration. The commercial rigid registration was on par with the control group (Figure 4b and Table 2). An example for the perfusion maps calculated with different motion correction algorithms can be seen in Figure 1b.

The ratings of the colour-coded perfusion maps showed statistically significant intergroup differences, as both the deformable registration using 80-kVp and DE-blended images differ significantly from the rigid registration as well as the deformable registration using VNC images and the control group (Table 3 part b).

On average, deformable registration using 80-kVp or DE-blended images achieved the best model fits, indicated by the lowest residuals (Figure 5a,b) and the highest R^2 values (Figure 5c,d), followed by the commercial non-rigid registration. The average R^2 values and residuals of the commercial rigid registration did not improve in comparison with the control group (Table 2).

Figure 4. Descriptive statistics of (a) ratings of the motion-corrected images and (b) ratings of the perfusion maps (4 = best, 1 = worst). Note that the lower quartile of the ratings of perfusion maps corrected with the commercial rigid registration is at 1.5, as the sum of rated perfusion maps is an even number. DE, dual energy; VNC, virtual non-contrast.



The intergroup differences in R^2 values and residuals were statistically significant for all comparisons for both the maximum slope and the Patlak model with few exceptions, as noted in Table 3 parts c and d.

Correlation between qualitative and quantitative evaluation

The correlation between the absolute values of the mean squares metric after motion correction and the average ratings of the motion-corrected images was moderate (−0.44). The correlation

Table 2. Average values ± standard deviation of the quantitative and qualitative evaluation

Evaluated scenario	Ratings motion-corrected images	Relative reductions in metric	Ratings perfusion maps	R^2 values maximum-slope	Residuals maximum-slope (HU)	R^2 values Patlak	Residuals Patlak
Commercial rigid	2.7 ± 1.0	0.6% ± 63.5%	2.4 ± 1.0	0.03 ± 0.13	526 ± 214	0.31 ± 0.14	1.33 ± 0.76
Commercial non-rigid	3.1 ± 0.9	29.2% ± 30.8%	2.8 ± 0.9	0.11 ± 0.07	350 ± 108	0.35 ± 0.16	0.95 ± 0.49
Deformable	3.3 ± 0.8	48.5% ± 23.4%	3.1 ± 0.9	0.18 ± 0.07	291 ± 71	0.41 ± 0.17	0.83 ± 0.47
Deformable dual-energy-blended	3.3 ± 0.8	45.6% ± 24.1%	3.0 ± 0.9	0.19 ± 0.06	290 ± 73	0.40 ± 0.19	0.84 ± 0.47
Deformable virtual non-contrast	2.9 ± 1.0	22.8% ± 27.4%	2.6 ± 0.9	0.12 ± 0.06	419 ± 103	0.33 ± 0.15	1.13 ± 0.49
Control	2.7 ± 1.1	0% ± 0%	2.5 ± 0.9	0.07 ± 0.06	509 ± 179	0.31 ± 0.15	1.32 ± 0.69

between the average R^2 values and residuals, and the average rating of the perfusion maps calculated with the maximum slope model was moderate (0.52/−0.31). The correlation between the average R^2 values and residuals, and the average perfusion maps calculated with the Patlak model was moderate to low (permeability, 0.29/−0.25; blood volume, 0.18/−0.20). Note that the negative signs are a result of the ascending order of the ratings and the descending order of the metric values and residuals.

DISCUSSION

The aim of this study was to provide a qualitative and quantitative evaluation of the effectiveness of different motion correction algorithms, as well as the use of DECT post processing, for the correction of breathing motion when calculating CT perfusion in abdominal organs.

The qualitative and quantitative evaluations show a clear benefit in using non-rigid registration algorithms, such as the commercial non-rigid algorithm or the self-developed deformable algorithm, for the correction of breathing motion when calculating CT perfusion in abdominal organs. Compared with the rigid motion-correction algorithm, there is a clear advantage in using non-rigid motion correction algorithms, which is in agreement with the findings of Chandler et al[9,10] for lung and liver tumours. Using the deformable registration presented here leads to an improvement in the spatial alignment of the ROI, as well as a better fit of the perfusion models and better quality of the resulting perfusion maps.

Compared with a state-of-the-art commercial non-rigid registration, the self-developed deformable algorithm yields better results, at the cost of higher computation times. While a computation time of up to several hours is a severe limitation unacceptable in clinical routine, it seems an acceptable limitation in view of clinical research, under the condition that the results are not needed immediately. When running the self-developed deformable algorithm for motion correction overnight, results are available for further clinical evaluation the morning after CT acquisition and initial segmentation of the reference image, if the workflow is adapted accordingly. Furthermore, none of the common strategies for the optimization and minimization of computation time such as the use of graphics processing unit computing or multilevel registration was employed in this article. These methods can be expected to provide considerable improvements in computation time by at least a factor of 10, permitting a reduction of execution times to the level of several minutes, that is, to a maximum of about 2–30 min per patient in cases of our study collective.[21]

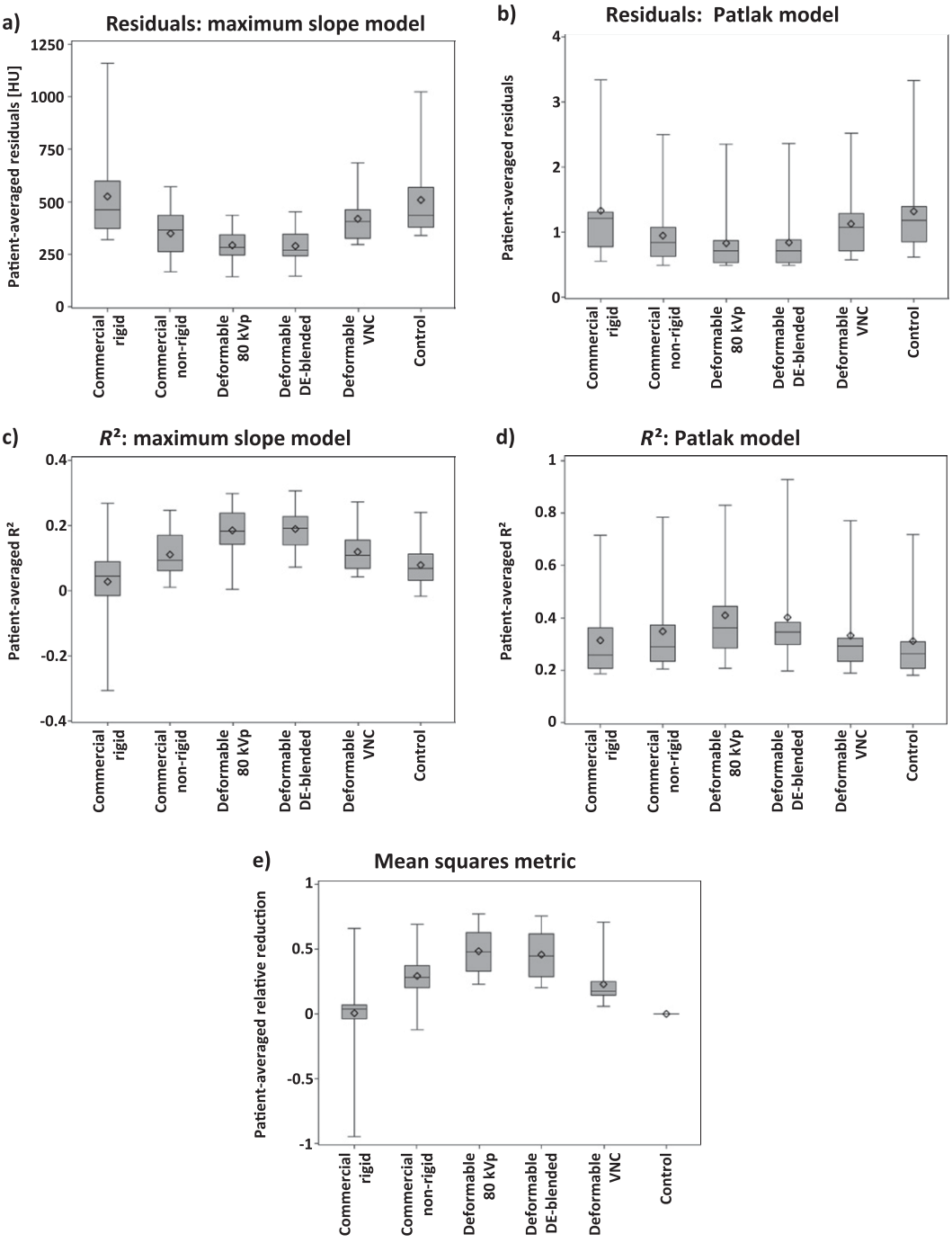
The use of DE image post-processing techniques does not provide any additional improvement of the registration. While the use of DE-blended images, which can be regarded as equivalent to a dose-equivalent single-energy 120-kVp image, for deformable registration achieved results on par with the use of 80-kVp images, the VNC images achieved significantly worse results.[13] This might be attributed to the higher noise level of the VNC images, limiting the quality of motion correction, which could be alleviated by using image filters.[22] Furthermore, the ITK implementation of the Demons deformable registration

Table 3. Statistically significant intergroup differences of (a) ratings of the motion-corrected images, (b) ratings of the perfusion maps, (c) R^2 values and residuals of the maximum slope model and (d) R^2 values and residuals of the Patlak model

Evaluated scenario	Commercial rigid	Commercial non-rigid	Deformable	Deformable DE-blended	Deformable VNC	Control
a) Ratings: motion-corrected images						
Commercial rigid		✓	✓	✓	✓	×
Commercial non-rigid			✓	✓	✓	✓
Deformable				×	✓	✓
Deformable DE-blended					✓	✓
Deformable VNC						✓
Control						
b) Ratings: perfusion maps						
Commercial rigid		✓	✓	✓	×	×
Commercial non-rigid			×	×	×	×
Deformable				×	✓	✓
Deformable DE-blended					✓	✓
Deformable VNC						×
Control						
c) R^2 values and residuals: maximum slope model						
Commercial rigid		✓	✓	✓	✓	✓
Commercial non-rigid			✓	✓	Res.	✓
Deformable				×	✓	✓
Deformable DE-blended					✓	✓
Deformable VNC						✓
Control						
d) R^2 values and residuals: Patlak model						
Commercial rigid		✓	✓	✓	✓	×
Commercial non-rigid			✓	✓	✓	✓
Deformable				×	✓	✓
Deformable DE-blended					✓	✓
Deformable VNC						✓
Control						

✓, significant; ×, not significant; DE, dual-energy; Res., only significant for residuals; VNC, virtual non-contrast.
Note that the lower half of the tables is omitted because of the inherent symmetry of the statistical tests.

Figure 5. Descriptive statistics of the quantitative evaluation: (a) residuals of the maximum slope model, (b) residuals of the Patlak model, (c) R^2 values of the maximum slope model, (d) R^2 values of the Patlak model, (e) relative reduction in metric values. Patient-wise averaging was performed to reduce the effect of different-sized regions of interest. Note the different units of the residuals, as the residual of the maximum slope model is calculated from CT values, while the residual of the Patlak model is calculated from the Patlak graph. The *y*-axes have been scaled differently in order to allow for visualization of subtle differences. DE, dual-energy; VNC, virtual non-contrast.



algorithm already includes a histogram normalization filter, which reduces the problem of intensity differences between images of different acquisition times.

While the descriptive statistics show a clear benefit in using the commercial non-rigid registration in comparison with the deformable registration using VNC images or the control group, the differences in the ratings of the perfusion maps are not statistically significant. This discrepancy indicates that the patient collective of 12 patients is too small to achieve the statistical power necessary for this comparison. However, the evaluation was limited by the high number of readings needed for the

assessment of the motion-corrected images, as each radiologist had to rate the spatial alignment of nearly 2500 images. Therefore, future studies might use quantitative measures for the evaluation of spatial alignment of motion-corrected images such as the metric values obtained during the optimization process, the percentage overlap or distance of centre of mass.[10]

The moderate correlation between the qualitative and quantitative evaluation indicates that information available to the software during motion correction and calculation of the perfusion maps might provide a sufficient assessment of quality. Furthermore, descriptive statistics obtained from the quantitative evaluation are in agreement with the ratings performed by experienced radiologists. However, the correlation was noticeably lower for the perfusion maps calculated with the Patlak model. The most probable reason for this discrepancy is the higher distance of the outliers in comparison with the quartiles for both average R^2 and residuals observed with the Patlak model for all motion correction algorithms, as these distances are not adequately represented by a rating on a four-point scale. Furthermore, the low correlation might indicate that the parameters provided by the Patlak model, permeability and blood volume are harder to interpret than blood flow. A more sophisticated evaluation with a larger patient collective could provide additional insights into this issue.

As a technical limitation of the CT scanner, the acquisition of DECT perfusion measurements is limited to a collimation of 32×0.6 mm (Table 1). This is much smaller than the full size of the scanned organs and tissues, and comparable abdominal perfusion studies employ a larger z-coverage.[5,7] As the impact of the breathing motion can be expected to be much more prominent when using a small z-coverage, the motion correction might be of less importance when using whole organ coverage. Furthermore, the motion correction was performed without any manual intervention to exclude data points affected by particularly heavy breathing motion, as would be usual in a clinical setting.[3] However, while these limitations can be expected to affect the quality of the resulting perfusion maps, the effect should be equal for all of the evaluated algorithms, and inter-patient effects were accounted for by using a GLM.

Another limitation of our work is the fact that the commercial rigid motion correction algorithm had to be recreated in the software suite developed in-house, as technical limitations of the manufacturer software prevent the export of motion-corrected image data. While the specifications of the vendor on functionality of the algorithm are extensive, and the visual comparison showed comparable results, this could heavily influence the evaluation. However, the results on the performance of the rigid algorithm are consistent with earlier studies performed on the manufacturer workstation.[23]

Additionally, only a small number of different motion correction algorithms were evaluated in this study, while a great number of possible algorithms have been proposed in the literature and more extensive evaluations have been performed for different applications.[6,8] Therefore, this evaluation should not be regarded as the only possible solution to the problem of motion correction for abdominal CT perfusion. Further evaluation seems necessary to fully explore the effects of different types of motion correction algorithms and their possible combination with image post-processing and acquisition techniques for the correction of breathing motion in abdominal organs for CT perfusion measurements.

While the qualitative and quantitative evaluation of the performance of different motion correction algorithms are in agreement, the evaluation is nonetheless limited, as no ground truth is available for perfusion measurements. As an improvement, future studies could evaluate the diagnostic value of the perfusion maps for the detection of tumours and other diseases, as often performed in studies for the evaluation of CT perfusion measurements.[4,7]

CONCLUSIONS

CT perfusion is a valuable tool for the detection of tumours, but breathing-induced motion of the abdominal organs can diminish the quality of the calculated perfusion maps. The use of the deformable motion correction algorithms presented here leads to significant improvements of spatial alignment of the image data as well as the quality of the resulting perfusion maps, whereas the use of DE post-processed images does not yield a significant benefit for the motion correction.

FUNDING

This research was supported by the German Research Foundation (DFG) within project R02, SFB/TRR 125 "Cognition-guided Surgery".

REFERENCES

1. Lee TY. Functional CT: physiological models. *Trends Biotechnol* 2002; **20**: S3–10.
2. Miles KA, Griffiths MR. Perfusion CT: a worthwhile enhancement? *Br J Radiol* 2003; **76**: 220–31.
3. Miles KA. Perfusion CT for the assessment of tumour vascularity: which protocol? *Br J Radiol* 2003; **76**: s36–42.
4. Klauss M, Stiller W, Pahn G, Fritz F, Kieser M, Werner J, et al. Dual-energy perfusion-CT of pancreatic adenocarcinoma. *Eur J Radiol* 2013; **82**: 208–14. doi: 10.1016/j.ejrad.2012.09.012
5. d'Assignies G, Couvelard A, Bahrami S, Vullierme MP, Hammel P, Hentic O, et al. Pancreatic endocrine tumors: tumor blood flow assessed with perfusion CT reflects angiogenesis and correlates with prognostic factors. *Radiology* 2009; **250**: 407–16. doi: 10.1148/radiol.2501080291
6. Crum WR, Hartkens T, Hill DL. Non-rigid image registration: theory and practice. *Br J Radiol* 2004; **77**: S140–53.
7. Kandel S, Kloeters C, Meyer H, Hein P, Hilbig A, Rogalla P. Whole-organ perfusion of the pancreas using dynamic volume CT in patients with primary pancreas carcinoma: acquisition technique, post-processing and initial results. *Eur Radiol* 2009; **19**: 2641–6. doi: 10.1007/s00330-009-1453-z

8. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 2009; **46**: 786–802. doi: 10.1016/j.neuroimage.2008.12.037
9. Chandler A, Wei W, Anderson EF, Herron DH, Ye Z, Ng CS. Validation of motion correction techniques for liver CT perfusion studies. *Br J Radiol* 2012; **85**: e514–22. doi: 10.1259/bjr/31999821
10. Chandler A, Wei W, Herron DH, Anderson EF, Johnson VE, Ng CS. Semiautomated motion correction of tumors in lung CT-perfusion studies. *Acad Radiol* 2011; **18**: 286–93. doi: 10.1016/j.acra.2010.10.008
11. Alvarez RE, Macovski A. Energy-selective reconstructions in X-ray computerized tomography. *Phys Med Biol* 1976; **21**: 733–44.
12. Johnson TR, Krauss B, Sedlmair M, Grasruck M, Bruder H, Morhard D, et al. Material differentiation by dual energy CT: initial experience. *Eur Radiol* 2007; **17**: 1510–17.
13. Behrendt FF, Schmidt B, Plumhans C, Keil S, Woodruff SG, Ackermann D, et al. Image fusion in dual energy computed tomography: effect on contrast enhancement, signal-to-noise ratio and image quality in computed tomography angiography. *Invest Radiol* 2006; **44**: 1–6. doi: 10.1097/RLI.0b013e31818c3d4b
14. Shrimpton P. *Assessment of patient dose in CT: European guidelines for multislice computed tomography funded by the European Commission 2004: contract number FIGMCT2000-20078-CT-TIP.* Luxembourg: European Commission; 2004.
15. itk.org. The Insight Segmentation and Registration Toolkit. [Cited 25 November 2014.] Available from: http://www.itk.org/
16. Brun R, Rademakers F. ROOT—an object oriented data analysis framework. *Nucl Instrum Meth A* 1997; **389**: 81–6.
17. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal* 1998; **2**: 243–60.
18. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. *Siam J Sci Comput* 1995; **16**: 1190–208.
19. Miles KA. Measurement of tissue perfusion by dynamic computed tomography. *Br J Radiol* 1991; **64**: 409–12.
20. Patlak CS, Blasberg RG. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. *J Cereb Blood Flow Metab* 1985; **5**: 584–90.
21. Muyan-Ozcelik P, Owens JD, Junyi X, Samant SS. Fast deformable registration on the GPU: a CUDA implementation of demons. In: *Proceedings of the 2008 international conference on computational sciences and its applications*. Los Alamitos, CA: IEEE Computer Society; 2008.
22. Kalender WA, Klotz E, Kostaridou L. An algorithm for noise suppression in dual energy CT material density images. *IEEE Trans Med Imaging* 1988; **7**: 218–24.
23. Skornitzke S, Fritz F, Klauss M, Pahn G, Grenacher L, Kauczor HU, et al. Evaluation of the effectiveness of rigid and deformable motion-correction algorithms with regard to CT-perfusion measurements in adominal organs affected by breathing motion. In: ECR 2014, Part B. Proceedings of the 2014 European Congress of Radiology; 6–10 March 2014; Vienna, Austria. Berlin, Germany: Springer 2014.