

Validation of nonrigid registration in pretreatment and follow-up PET/CT scans for quantification of tumor residue in lung cancer patients

Jolanda Spijkerman,¹ Davide Fontanarosa,^{1,2a} Marco Das,³
Wouter van Elmpt¹

*Department of Radiation Oncology (MAASTRO),¹ GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands; Oncology Solutions,² Philips Research, Eindhoven, The Netherlands; Department of Radiology,³ GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands
davide.fontanarosa@maastro.nl*

Received 10 December, 2013; accepted 20 March, 2014

Nonrigid registrations of pre- and postradiotherapy (RT) PET/CT scans of NSCLC patients were performed with different algorithms and validated tracking internal landmarks. Dice overlap ratios (DR) of high FDG-uptake areas in registered PET/CT scans were then calculated to study patterns of relapse. For 22 patients, pre- and post-RT PET/CT scans were registered first rigidly and then nonrigidly. For three patients, two types (based on Demons or Morphons) of nonrigid registration algorithms each with four different parameter settings were applied and assessed using landmark validation. The two best performing methods were tested on all patients, who were then classified into three groups: large (Group 1), minor (Group 2) or insufficient improvement (Group 3) of registration accuracy. For Group 1 and 2, DRs between high FDG-uptake areas in pre- and post-RT PET scans were determined. Distances between corresponding landmarks on deformed pre-RT and post-RT scans decreased for all registration methods. Differences between Demons and Morphons methods were smaller than 1 mm. For Group 1, landmark distance decreased from 9.5 ± 2.1 mm to 3.8 ± 1.2 mm (mean \pm 1 SD, $p < 0.001$), and for Group 3 from 13.6 ± 3.2 mm to 8.0 ± 2.2 mm ($p = 0.025$). No significant change was observed for Group 2 where distances decreased from 5.6 ± 1.3 mm to 4.5 ± 1.1 mm ($p = 0.093$). DRs of high FDG-uptake areas improved significantly after nonrigid registration for most patients in Group 1. Landmark validation of nonrigid registration methods for follow-up CT imaging in NSCLC is necessary. Nonrigid registration significantly improves matching between pre- and post-RT CT scans for a subset of patients, although not in all patients. Hence, the quality of the registration needs to be assessed for each patient individually. Successful nonrigid registration increased the overlap between pre- and post-RT high FDG-uptake regions.

PACS number: 87.57.Q-, 87.57.C-, 87.57.N-, 87.57.-s, 87.55.-x, 87.55.D-, 87.55.dh, 87.57.uk, 87.57.nj

Key words: PET, deformable algorithm, pattern of relapse

I. INTRODUCTION

Although prognosis of lung cancer patients has improved over the last decade, it remains one of the most lethal solid tumors. The five-year survival is approximately 20% in patients with locally advanced non-small cell lung cancer (NSCLC) treated with radiotherapy (RT), but, in

^a Corresponding author: Davide Fontanarosa, Maastricht Clinic, Dr Tanslaan 12, 6229ET Maastricht, the Netherlands; phone: +31 (0) 6 11534106; fax: +31 88 4455667; email: davide.fontanarosa@maastro.nl

the majority of patients, local relapses are observed after treatment.⁽¹⁻³⁾ In order to improve treatment in future and design strategies to prevent local relapses, follow-up scans of lung cancer patients are used to accurately quantify tumor treatment response patterns.

To find patterns of relapse, rigid registration between pretreatment (pre-RT) and follow-up posttreatment (post-RT) positron emission tomography (PET)/computerized tomography (CT) scans was used mainly.⁽⁴⁻⁶⁾ Previous work that only used rigid registrations showed that the high FDG-uptake areas prior to treatment largely corresponded to tumor residue areas.⁽⁷⁾ However, patients with significant anatomy changes during RT (e.g., large tumor shrinkage or deformation, weight loss or change of body position during scanning) had to be excluded. In this work, we propose a solution that uses deformable (nonrigid) registration to correct for such anatomical changes and allows investigation of the treatment response on a voxel-to-voxel level. Deformable registration was previously used to register CT scans in different respiratory phases for tumor tracking,^(8,9) or for dose accumulation.^(10,11) However, to our knowledge this is the first study that investigates deformable registration of follow-up CT scans of lung cancer patients taken more than four months after pre-RT scans.

With deformable registration, it is paramount to validate the deformation fields. Validation can be performed by defining landmarks on visible features inside the thorax (e.g., bifurcating vessels, calcifications, distinct anatomical shapes) in order to quantify the accuracy of the deformation fields.⁽¹²⁻¹⁴⁾

The verified registration fields can then be used to investigate the relation between high [18F]-fluorodeoxyglucose (FDG)-uptake areas in the tumor before the treatment and tumor residue three months after treatment. FDG-uptake could serve as an indication of treatment-resistant regions inside the tumor.⁽⁶⁾ In this research, the verified deformable registration fields calculated on the CT scans were applied to the PET scans. Since deformable registration could compensate for the anatomical changes that occur in the months between the scans, we hypothesized that the overlap between high FDG-uptake areas is increased compared to rigid registration.

To test this hypothesis, we have two aims for this study: first, to validate deformable image registration between pre- and post-RT PET/CT image datasets of NSCLC patients; second, if accurate deformable registration is possible, to test the hypothesis that deformable registration techniques yield a better match than rigid transformation. For this, we evaluated Dice overlap ratios (DR) of tumor residue in the post-RT scans with high FDG-uptake areas in the pre-RT PET scan. This research is a prerequisite for calculating voxel-by-voxel correspondence used in evaluating dose-painting strategies.

II. MATERIALS AND METHODS

A. Patient data

Twenty-two patients (six female and 16 male) were included, with inoperable NSCLC, UICC stage I–III, treated with radical RT alone or with sequential chemo-RT, according to the institutional protocol.⁽⁶⁾ The datasets are described comprehensively elsewhere.⁽⁶⁾ No treatment was given to any of the patients between the end of RT and the post-RT scan. A PET/CT scanner (Siemens Biograph 40, Knoxville, TN) was used to image the patients, in treatment position. The CT part of the scan was acquired using 4D CT, the PET was acquired using 3D PET. The scans used in the analysis were in the midventilation phase of the 4D CT, the same phase was also used for attenuation correction of the PET scan.⁽¹⁵⁾ Two PET/CT scans were acquired: one before RT treatment and one approximately three months after treatment. The CT scans used in this study were noncontrast-enhanced. Resolution of CT scans ranged from $0.7 \times 0.7 \times 3 \text{ mm}^3$ to $1 \times 1 \times 5 \text{ mm}^3$. For all PET scans the voxel size was $5.3 \times 5.3 \times 5 \text{ mm}^3$.

B. Deformable registration and extensive validation

If a large initial mismatch between pre- and post-RT CT scans was present, the scans were initially matched using automatic rigid registration. This automatic registration mainly matches the patient's contour in both scans. However, due to weight loss, for example, the size and shape may change. Therefore, for all patients, a manual rigid registration was applied to optimize the match between lung contours. A volume of interest (VOI) was created as the smallest box encompassing the PTV in the pretreatment scan with a further expansion of 20 mm. Since the computation time of nonrigid registration fields strongly depends on the image size, deformable registration was only calculated inside this VOI. In order to get a reliable registration field, the cropped pre- and post-RT scans should both contain the same anatomical structures. Twenty millimeters is the minimal suitable margin for this purpose, according to our experience. Both the rigid and deformable registrations were performed using in-house-developed software.⁽¹⁶⁾

For three randomly selected patients, eight deformable registration fields were calculated to deform the pre-RT CT scan to match the post-RT CT scan. These fields were based on the Demons or Morphons methods, with 10 or 20 iterations per scale (eight resolution scales, each scale is an increase of factor 2, up to the full resolution⁽¹⁰⁾). A weighted sum accumulation (e.g., every iteration updates the previous calculation) or a diffeomorphic (e.g., invertible and smooth) accumulation of the deformation field was performed with a regularization filter (e.g., Gaussian smoothing of the calculated deformation vectors) of 1.5 times the voxel size. The final calculation scale was the full CT resolution. Demons is a fast algorithm based on image intensity differences. Morphons is based on intensity phase differences (e.g., local intensity gradients in the image). Details on the exact parameter setting for these two types of registration methods and four different choices of parameter settings are shown in Table 1. More details of the methods are described in the literature.^(17,18) The landmarks in the deformed CT datasets of these three patients were extensively investigated to determine which deformation algorithms performed best. Together with an experienced radiologist, we annotated both the pre-RT and post-RT CT scan using approximately 30 landmarks/fiducials per patient. This included approximately ten landmarks inside or directly surrounding the primary lung tumor. These landmarks were selected based on visible features inside the thorax (e.g., bifurcating vessels, calcifications and distinct anatomical shapes). In Fig. 1, one landmark is reported as an example. As a measure of accuracy, the landmarks were tracked in the deformed scans, and absolute differences between deformed and nondeformed annotated landmark positions were calculated and compared for the various algorithms. It is important to notice that this validation method determines the accuracy of the registration field for a particular patient, not the accuracy of the deformation algorithms, as such.

Based on accuracy and computation times obtained with these three patients, we selected two parameter sets for deformable registration field calculation for the remaining 19 patients, one parameter set optimized for the Demons algorithm and one for the Morphons. For both algorithms, ten iterations per resolution scale and weighted sum accumulation of the update field were used. For these additional 19 patients, the accuracy of the two deformable registration fields was checked by annotating and tracking ten landmarks per patient.

Validation of rigid and deformable registration was quantified using the average absolute distance (3D vector) between the landmarks in the (deformed) pre- and post-RT scans. The patients were then classified into three groups based on the following criteria:

- Group 1: Large improvement of the match between the pre- and post-RT scans after deformable registration, compared to rigid registration; mean distance after deformable registration is less than half of the distance after rigid registration; mean distance after deformable registration is smaller than 6 mm.
- Group 2: No or minor improvement of the match between the pre- and post RT scans after deformable registration; mean distance after deformable registration is more than half of

the distance after rigid registration; mean distance after rigid registration is smaller than 7.5 mm; mean distance after deformable registration is smaller than 6 mm.

- Group 3: Inadequate match after rigid and deformable registration; mean distance after rigid registration is larger than 7.5 mm; mean distance after deformable registration is larger than 6 mm.

The minimal distance of 6 mm is comparable to the PET resolution ($5.3 \times 5.3 \times 5 \text{ mm}^3$), whereas the maximal distance was chosen as approximately 1.5 times the PET resolution as an upper limit of the necessary registration accuracy.

TABLE 1. Absolute mean landmark distances and landmark distance range (smallest to largest distance), for the three patients after rigid and deformable registration. The applied algorithm features are described as Algorithm (rigid, based on Demons or based on Morphons), weighted sum/diffeomorphic field accumulation (W/D), and number of iterations.

	Algorithm	W/D	Iterations	Mean \pm SD Entire VOI [mm] (range)	Mean \pm SD Tumour [mm] (range)	Calculation Time
Patient 1 31 Landmarks, 10 near tumour	Rigid			5.7 \pm 2.1 (1.0 - 10.2)	5.6 \pm 2.1 (2.8 - 8.5)	
	Demons	w	10	4.2 \pm 2.1 (1.0 - 7.6)	4.3 \pm 2.1 (1.0 - 6.6)	10min
	Demons	D	10	3.6 \pm 2.5 (1.0 - 10.9)	3.2 \pm 2.2 (1.0 - 7.6)	43min
	Demons	w	20	3.8 \pm 2.2 (1.0 - 10.2)	3.6 \pm 2.1 (1.0 - 6.6)	1h 12 min
	Demons	D	20	3.8 \pm 2.3 (1.0 - 10.1)	3.4 \pm 2.4 (1.0 - 7.6)	2h 34min
	Morphons	w	10	3.3 \pm 2.7 (0 - 10.6)	3.2 \pm 1.9 (1.0 - 5.4)	18min
	Morphons	D	10	3.6 \pm 3.3 (0 - 15.3)	3.0 \pm 1.9 (1.0 - 5.4)	1h 3min
	Morphons	w	20	3.5 \pm 2.7 (0 - 10.6)	3.1 \pm 1.9 (1.0 - 5.4)	2h 53min
	Morphons	D	20	3.2 \pm 2.5 (0 - 10.2)	2.8 \pm 1.8 (1.0 - 5.2)	4h 17 min
Patient 2 30 landmarks, 12 near tumour	Rigid			8.7 \pm 4.2 (3.0 - 17.2)	9.7 \pm 3.9 (3.0 - 16.1)	
	Demons	w	10	7.2 \pm 3.1 (1.0 - 11.5)	6.8 \pm 3.2 (3.0 - 11.5)	11min
	Demons	D	10	6.5 \pm 3.4 (1.0 - 12.6)	6.0 \pm 4.0 (1.0 - 12.6)	60min
	Demons	w	20	7.2 \pm 3.2 (0 - 11.5)	6.9 \pm 3.1 (3.0 - 11.5)	1h 20min
	Demons	D	20	6.3 \pm 3.6 (0 - 12.6)	5.9 \pm 4.2 (1.0 - 12.6)	3h 9min
	Morphons	w	10	4.4 \pm 2.6 (0 - 9.5)	3.0 \pm 2.2 (0 - 7.9)	27min
	Morphons	D	10	5.6 \pm 3.2 (0 - 10.7)	4.9 \pm 3.4 (0 - 10.7)	51min
	Morphons	w	20	4.5 \pm 2.4 (0 - 8.6)	3.1 \pm 2.0 (0 - 7.4)	1h 50min
	Morphons	D	20	3.7 \pm 1.9 (0 - 6.9)	2.1 \pm 1.9 (0 - 5.5)	3h 53min
Patient 3 32 landmarks, 11 near tumour	Rigid			8.9 \pm 4.2 (2.2 - 17.1)	10.2 \pm 4.5 (2.2 - 17.1)	
	Demons	w	10	6.9 \pm 4.7 (0 - 18.0)	10.1 \pm 5.5 (1.4 - 18.0)	10min
	Demons	D	10	6.4 \pm 4.4 (0 - 15.3)	9.3 \pm 5.3 (1.4 - 15.3)	45min
	Demons	w	20	6.8 \pm 4.6 (0 - 17.5)	10.4 \pm 5.5 (1.4 - 17.5)	2h 6min
	Demons	D	20	6.2 \pm 4.4 (0 - 15.3)	9.4 \pm 5.0 (1.4 - 15.3)	3h 2min
	Morphons	w	10	6.8 \pm 4.8 (0 - 16.9)	9.6 \pm 5.5 (1.4 - 16.9)	51min
	Morphons	D	10	6.4 \pm 4.3 (0 - 16.9)	9.1 \pm 5.8 (0 - 16.9)	1h 30min
	Morphons	w	20	7.0 \pm 4.7 (1.0 - 16.9)	10.1 \pm 5.4 (1.4 - 16.9)	2h 21min
	Morphons	D	20	6.3 \pm 4.3 (0 - 16.9)	9.1 \pm 5.7 (0 - 16.9)	4h 26min

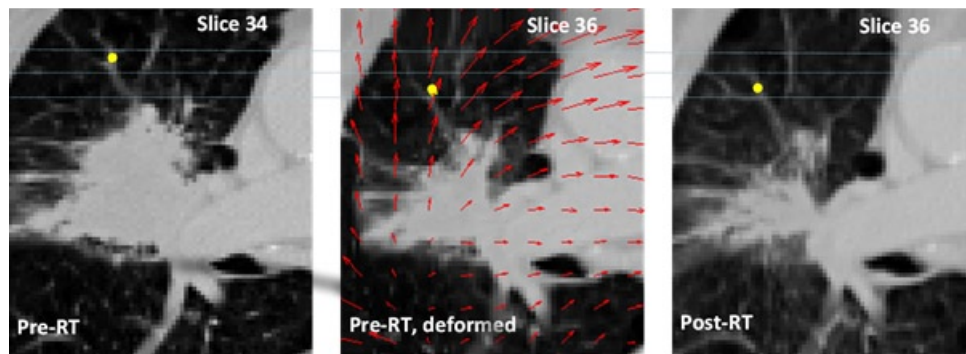


FIG. 1. Representative image for a patient in Group 1 showing a landmark in the pre-RT scan (left image) and post-RT CT scan (right image, after rigid registration), the deformed pre-RT CT scan is shown in the middle. Notice that there is also a difference in craniocaudal distance indicated by the slice numbers. Originally the landmarks in pre- and post-RT scan were two slices apart, but after deformation they are located on the same slice. The deformation field is indicated by red arrows in the deformed scan. Note that the arrows point in the opposite direction of the displacement vectors. The landmark distance was reduced from 14 mm to 3 mm after deformable registration.

C. PET scan analysis

Analysis of residual uptake patterns measured on the FDG-PET scans was performed for the patients with a good match after registration (Group 1 and 2); Group 3 was excluded from further analysis. To allow comparison with previous work,⁽⁵⁾ the contours corresponding to 34%, 40%, 50%, 60%, and 70% of SUV_{max} high-uptake values in the pre-RT scans were outlined, based on the corresponding thresholds. The contours of 70%, 80%, and 90% of SUV_{max} high-uptake values were outlined, instead, on the post-RT scans. Different thresholds were used for the pre- and post-RT scans, because SUV values in tumor residue in the post-RT PET scans are generally lower than SUV values in the tumor in the pre-RT scan. The registration fields resulting from rigid and deformable registration of CT scans were applied to the contours in the nonregistered scans. Because the voxel values can change due to deformation and interpolation effects, the PET image was not deformed, but only the thresholded contours. Then, the DR between the contours in the (deformed) pretreatment and posttreatment thresholded PET volumes were determined according to the following definition:⁽¹⁹⁾

$$DR = \frac{2 \times V_{A \cap B}}{V_A + V_B} \quad (1)$$

where V_A and V_B represent the volumes of contours A and B, and $V_{A \cap B}$ represents their intersection. The DRs between volumes in post-RT scans and rigidly registered pre-RT scans were first evaluated. Then, also the DRs between volumes in the post-RT and the deformed pre-RT scans were calculated.

D. Statistical analysis

Descriptive statistics of deformations were given in mean values ± 1 standard deviation (SD). Paired *t*-tests were performed to compare mean landmark distances before and after deformable registration. Wilcoxon signed-rank tests were performed to compare the DR between the contours in the post-RT PET scan and the pre-RT (deformed) PET scan before and after deformable registration. P-values smaller than 0.05 were assumed to be statistically significant.

III. RESULTS

A. Validation of registration fields

For the three patients with the 30 landmarks defined, the absolute distances between landmarks in the deformed pre- and post-RT scans were calculated and summarized in Table 1 for rigid registration and for the eight deformable registration fields. Calculation times of the deformable registration fields ranged from 10 min to 4.5 hrs, depending on the algorithm and size of the VOI. An example of a single landmark selected in the pre- and post-RT CT scans and tracked in the deformed pre-RT scan for a patient in Group 1, is shown in Fig. 1. In the deformed pre-RT CT scan the deformation field is indicated using arrows.

In five out of six cases, the mean landmark distances were smallest for the Morphons algorithm, with 20 iterations per scale and diffeomorphic field accumulation. But it also had the longest computation times, approximately 4.5 hrs per patient. Differences in mean landmark distances within the four variants of the Demons and Morphons algorithm were small, ranging between 0.4 mm and 0.9 mm, which is comparable to the in-plane resolution of the CT scan. Although the landmark distances decreased between the best performing algorithm (either Demons or Morphons with 20 iterations per scale and diffeomorphic field accumulation) and the fastest Demons/Morphons algorithm (ten iterations per scale, weighted sum accumulation), this decrease was not statistically significant. Therefore, the fastest Demons variant and the fastest Morphons variant were selected for the subsequent analysis of all patients.

B. Selection of best registration fields

For one of the 22 patients, deformable registration was not successful due to atelectasis that was not present in the pre-RT scan and appeared on the post-RT scan. In another patient, pneumonitis occurred at the location of the GTV in the post-RT scan. These two patients were excluded from the analysis.

For the remaining 20 patients, the best deformable registration algorithm was determined based on the accuracy of the landmark analysis. Subsequently the patients were classified into 3 groups, as previously described. The mean absolute landmark distances for the patients in these groups are summarized in Table 2. For Group 1 and 3, the landmark distances decreased significantly after deformable registration ($p < 0.001$ and $p = 0.025$, respectively). The Morphons algorithm yielded the best results for six out of seven patients in Group 1, for three out of seven patients in Group 2, and for four out of six patients in Group 3.

TABLE 2. Mean absolute landmark distances for the three groups after rigid and after deformable registration. Mean \pm SD of the mean absolute landmark distances of the patients in each group are shown after rigid, Demons, and Morphons registration. Also the numbers for the best deformable registration for each patient are shown. The p -value results from comparing the absolute landmark distances after rigid registration to the distances after the best deformable registration.

	<i>N</i>	<i>Rigid</i> (mean \pm SD)	<i>Demons</i> (mean \pm SD)	<i>Morphons</i> (mean \pm SD)	<i>Best Deformable</i> <i>Registration</i> (mean \pm SD)	<i>p-value</i> <i>Rigid vs. Best</i>
Group 1	7	9.5 \pm 2.1 mm	5.4 \pm 1.7 mm	4.4 \pm 1.8 mm	3.8 \pm 1.2 mm	<0.001
Group 2	7	5.6 \pm 1.3 mm	4.9 \pm 1.1 mm	4.8 \pm 1.5 mm	4.5 \pm 1.1 mm	0.093
Group 3	6	13.6 \pm 3.2 mm	8.5 \pm 2.6 mm	10.2 \pm 2.7 mm	8.0 \pm 2.2 mm	0.025

N = number of patients per group.

C. PET Dice overlap ratios

For six of the seven patients in Group 1, large overlap between pre- and post-RT high FDG-uptake areas occurred. In Fig. 2, the (deformed) PET/CT scans of two patients in Group 1 are shown: the (deformed) 60% of SUV_{max} contours in the (deformed) pre-RT scan are compared

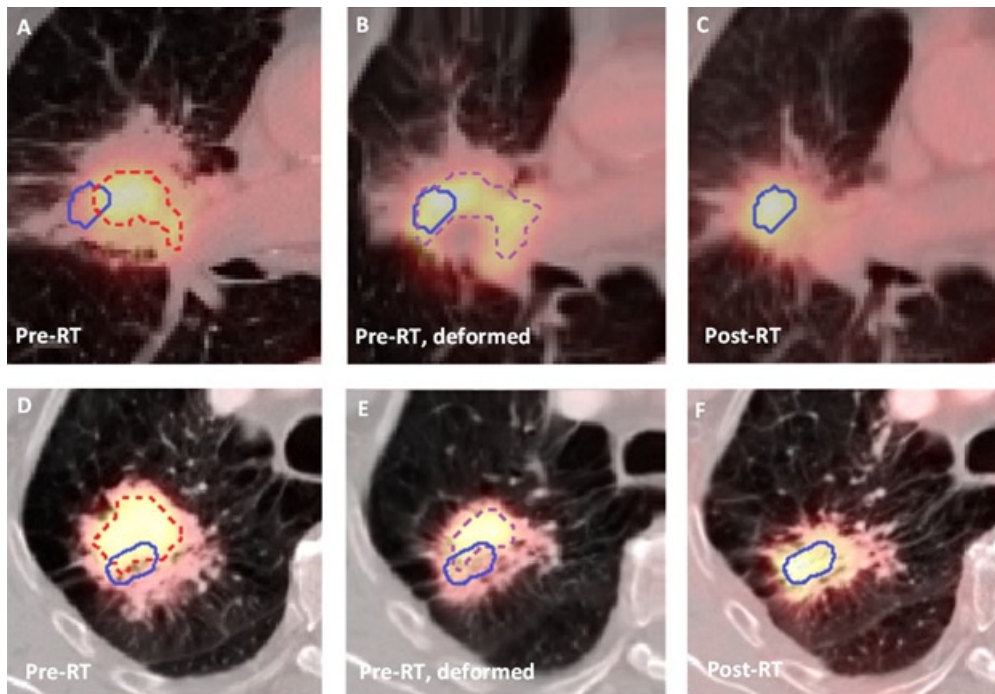


FIG. 2. (Top) FDG-PET/CT images of a patient in Group 1 with high overlap fractions: pre-RT scan (a), deformed pre-RT scan (b), post-RT scan (c). The 60% of SUV_{max} contour in the pre-RT scan (red, dashed), the deformed 60% of SUV_{max} contour in the deformed pre-RT scan (purple, dashed), and the 70% of SUV_{max} in the post-RT scan (blue) are depicted. (Bottom) FDG-PET/CT images of a patient in Group 1 with low overlap fractions: pre-RT scan (d), deformed pre-RT scan (e), rigidly registered post-RT scan (f). The 60% of SUV_{max} contour in the pre-RT scan (red, dashed), the deformed 60% of SUV_{max} contour in the deformed pre-RT scan (purple, dashed), and the 70% of SUV_{max} in the post-RT scan (blue) are depicted.

to the 70% of SUV_{max} contour in the post-RT scans. For the first patient (Figs. 2(a) and (c)) the tumor shape in the deformed scan visually better resembles the shape after RT. Also, in this slice, the 70% of SUV_{max} contour in the post-RT scan is almost entirely located within the deformed 60% of SUV_{max} contour, as opposite to the nondeformed pre-RT contour. After deformation of the pre-RT contour, the DR between the described contours for this patient increased from 6% to 21%. For the second patient (Figs. 2(d) and (f)), visual evaluation of the scans shows that the tumor's shape in the deformed pre-RT scan is quite similar to the post-RT scan. However, the 70% of SUV_{max} contour in the post-RT scan appears to have low overlap with the 60% of SUV_{max} contour in the deformed pre-RT scan. Comparing the scans, it seems that the hotspot in the pre-RT scan did not move to a new location, but a new hotspot has shown up in the post-RT PET-scan. The DR for this patient decreased from 19% to 12% after deformation of the pre-RT contour.

In Fig. 3, the DRs for the 70% and 90% of SUV_{max} contours in the post-RT scan are shown; DRs for the 80% of SUV_{max} threshold are similar (not shown). For Group 1, deformable registration of the pre-RT PET scan significantly increased almost all DRs compared to rigid registration; no significant difference was found for the 70% of SUV_{max} pre-RT contour and the 70%, 80%, and 90% of SUV_{max} post-RT contour, and for the 60% of SUV_{max} pre-RT contour and the 90% of SUV_{max} post-RT contour. The largest improvements were achieved for the 70% of SUV_{max} in the post-RT scan for this group. For Group 2, significant differences were only found for the DRs of the 34% of SUV_{max} pre-RT contours, and the 80% and 90% of SUV_{max} post-RT contour.

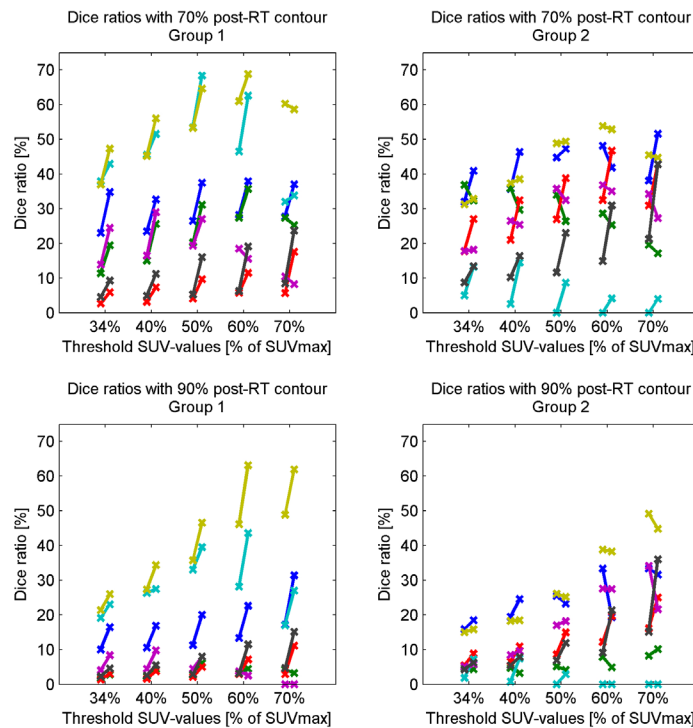


FIG. 3. Dice overlaps ratio of FDG uptake areas. For each patient (represented by different colors) paired values are plotted, before (left cross in the pair) and after (right cross) deformable registration for Group 1 and 2, connected by a line. Left, top to bottom: DRs of the (deformed) pre-RT contours with the 70% and 90% of SUV_{max} contours in the post-RT scan for Group 1. Right, top to bottom: DRs of the (deformed) pre-RT contours with the 70% and 90% of SUV_{max} contours in the post-RT scan for Group 2.

IV. DISCUSSION

Deformable registration fields need to be validated and the accuracy needs to be estimated before implementation in research or clinical practice.^(20,21) This work is, to our knowledge, the first study investigating the accuracy of image registration between pre-RT and follow-up imaging using scans that are four months apart. There are multiple deformation algorithms available nowadays that are utilized for several purposes: tumor tracking during the respiratory cycle,⁽⁸⁾ contour propagation for adaptive RT,⁽²²⁾ and dose accumulation.⁽¹⁰⁾ Validation datasets have been annotated with landmarks mainly for 4D CT imaging of the thorax,⁽²³⁾ but also other treatment sites and organs, are becoming available,⁽¹²⁾ but none are described for follow-up imaging. De Moor et al.⁽²⁴⁾ used a rigidity constraint to steer the deformation field calculation for PET to PET registrations in such a way that the tumor volume of interest remains constant for the purpose of response assessment. For CT to CT registration, this might also be an option to steer the deformation field calculation in regions with lower CT contrast; this needs, however, further investigation regarding the accuracy of such an approach. Another option is the use of synthetic datasets for validation purposes.⁽²⁰⁾ Once deformation algorithms are validated, they could also have potentially a role in dose assessment in a reirradiation setting,⁽²⁵⁾ for example in lung cancer, where large anatomical variations occur after treatment. The actual dose delivered to every organ can then serve as an indication of the dose level, and the actual high-dose region inside that organ that might then guide the treatment plan optimization for reirradiation.

In the first part of this work, we have chosen a landmark-based verification to extensively study three patients comparing eight deformable registration fields, based on Demons and Morphons, in order to estimate the accuracy of the various algorithms. Using the best performing algorithms, a significant improvement, compared to rigid registration alone, was achieved for 70% (14 out of 20) of the patients. It should be noted that for the remaining 30% of the patients, deformable registration did not significantly improve. This happened mainly for extremely large anatomy variations both for normal lung and tumor. In this case, modification of the deformation algorithms to take missing tissue into account might be necessary.⁽²⁶⁾ Also, for about two-thirds of the patients, the Morphons algorithm showed smallest deviations, while for one-third of them, it was Demons. Since we could not identify any causes why for some patients Morphons and for others Demons was the best, it is still necessary to perform individualized validation in order to determine which algorithm is the most accurate for each individual patient.

Although specific nonrigid registration algorithms exist that incorporate tissue properties (Al-Mayah et al.⁽²⁷⁾), the frequently used (mathematical) nonrigid registration algorithms do not model physical properties of organs or tumors. The algorithm will search for matching intensity values inside images, but could also try to match parts of the tumor that were in regression and started growing again. The deformation vector from the original tumor voxel to a newly grown tissue might look correct on an imaging level, but will represent different biological tissues. More frequent follow-up imaging and characterization of tumor regression patterns are needed to tackle these issues.

To perform response assessment of treatment on a local or voxel level, multiple strategies can be used. One of these techniques models the voxel control probability and could benefit from better image registration techniques.^(28,29) However, only few of the described methods have validated their registration accuracy prior to building their predictive models. In our study, the DR between pre- and post-RT FDG-uptake increased after deformable registration. As expected for Group 2, the deformed pre-RT CT scan showed no significant reduction in average landmark distance compared to rigid registration and, hence, the DR showed no improvement. This confirms our hypothesis that a successful deformable registration increases the overlap ratio between high FDG-uptake areas in pre- and post-RT PET scans. This further implies that FDG-uptake is a suitable target for dose-boosting/painting as an indicator of treatment-resistant regions inside the tumor.⁽³⁰⁾ Since image deformation and subsequent interpolation of PET scans may change the actual SUV values, it was decided to apply the deformation fields to the contours instead of the image datasets. In this study, PET contours were delineated using thresholding, but other techniques, like gradient-based segmentation,⁽³¹⁾ are available. This technique may yield more robust segmentation of the PET values and could be used in a future research.

The proposed technique is also directly applicable to other tracers. Bowen et al.⁽³²⁾ showed a tri-modality investigation linking proliferation and hypoxia PET to residual FDG-uptake. Although they limited the registration process to an affine registration, it shows the possibility of finding imaging and dose response relationships between multiple modalities. Hypoxia is another factor contributing to treatment resistance⁽³³⁾ and does not necessarily coincide with metabolic active areas. Therefore, the tumor might recur or have a residue in the hypoxic area, which cannot be predicted from the FDG-PET scan. Combining hypoxia and FDG-PET imaging could give more insight in the areas of treatment resistance.⁽²⁸⁾ The presented methodology could then directly be applied to pre-RT hypoxia PET imaging and the assessment of treatment recurrence.

V. CONCLUSIONS

We implemented a clinical validation framework for individualized voxel-to-voxel deformable registrations between pre- and post-RT PET/CT scan of lung cancer patients. Deformable

registration significantly improved registration quality between pre- and post-RT CT scans, but the quality of the registration needs to be evaluated for all patients individually. DR of pre- and post-RT high FDG-uptake regions increased using deformable registration, depending on the registration quality. This technique allows more precise assessment in follow-up imaging of the initial location where tumor relapses occur.

ACKNOWLEDGMENTS

This study was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), Project AIRFORCE No. 03O-103. We would like to thank Georgy Shakirin for initial discussions on the study design. One of the authors (W.v.E.) would like to acknowledge funding (KWF MAC 2011-4970) from the Dutch Cancer Society.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA: Cancer J Clin.* 2012;62(13):10–29.
2. Betticher DC, Hsu Schmitz SF, Totsch M, et al. Prognostic factors affecting long-term outcomes in patients with resected stage IIIA pN2 non-small-cell lung cancer: 5-year follow-up of a phase II study. *Br J Cancer.* 2006;94(8):1099–106.
3. vanMeerbeek JP, Kramer GW, vanSchil PE, et al. Randomized controlled trial of resection versus radiotherapy after induction chemotherapy in stage IIIA-N2 non-small-cell lung cancer. *J Nat Cancer Inst.* 2007;99(6):442–50.
4. Aerts HJ, Bosmans G, van Baardwijk AA, et al. Stability of 18F-deoxyglucose uptake locations within tumor during radiotherapy for NSCLC: a prospective study. *Int J Radiat Oncol Biol Phys.* 2008;71(5):1402–07.
5. Aerts HJ, Bussink J, Oyen WJ, et al. Identification of residual metabolic-active areas within NSCLC tumours using a pre-radiotherapy FDG-PET-CT scan: a prospective validation. *Lung Cancer.* 2012;75(1):73–76.
6. Aerts HJ, van Baardwijk AA, Petit SF, et al. Identification of residual metabolic-active areas within individual NSCLC tumours using a pre-radiotherapy (18)Fluorodeoxyglucose-PET-CT scan. *Radiother Oncol.* 2009;91(3):386–92.
7. Abramuk A, Tokalov S, Zophel K, et al. Is pre-therapeutic FDG-PET/CT capable to detect high risk tumor subvolumes responsible for local failure in non-small cell lung cancer? *Radiother Oncol.* 2009;91(3):399–404.
8. Orban deXivry J, Janssens G, Bosmans G, et al. Tumour delineation and cumulative dose computation in radiotherapy based on deformable registration of respiratory correlated CT images of lung cancer patients. *Radiother Oncol.* 2007;85(2):232–38.
9. Kadoya N, Fujita Y, Katsuta Y, et al. Evaluation of various deformable image registration algorithms for thoracic images. *J Radiat Res.* 2013;56(1):175–82.
10. Janssens G, de Xivry JO, Fekkes S, et al. Evaluation of nonrigid registration models for interfraction dose accumulation in radiotherapy. *Med Phys.* 2009;36(9):4268–76.
11. Li M, Castillo E, Zheng XL, et al. Modeling lung deformation: a combined deformable image registration method with spatially varying Young's modulus estimates. *Med Phys.* 2013;40(3):081902.
12. Brock KK. Results of a multi-institution deformable registration accuracy study (MIDRAS). *Int J Radiat Oncol Biol Phys.* 2010;76(2):583–96.
13. Samavati N, McGrath DM, Lee J, et al. Biomechanical model-based deformable registration of MRI and histopathology for clinical prostatectomy. *J Pathol Inform.* 2011;2:S10.
14. Yang D, Lu W, Low DA, Deasy JO, Hope AJ, El Naqa I. 4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling. *Med Phys.* 2008;35(10):4577–90.
15. Rosario T, Ollers MC, Bosmans G, De Ruyscher D, Lambin P, Dekker A. Phased versus midventilation attenuation-corrected respiration-correlated PET for patients with non-small cell lung cancer. *J Nucl Med Technol.* 2009;37(4):208–14.
16. Janssens G. Registration models for tracking organs and tumors in highly deformable anatomies: applications to radiotherapy [PhD thesis]. Belgium: Université Catholique de Louvain (UCL); 2010.
17. Nithianathan S, Schafer S, Uneri A, et al. Demons deformable registration of CT and cone-beam CT using an iterative intensity matching approach. *Med Phys.* 2011;38(4):1785–98.
18. Wrangsjö A, Pettersson J, Knutsson H. Non-rigid registration using morphons. *Image Analysis.* 2005;3540:501–10.
19. Yang X and Fei B. A multiscale and multiblock fuzzy C-means classification method for brain MR images. *Med Phys.* 2011;38(6):2879–91.
20. Nie K, Chuang C, Kirby N, Braunstein S, Pouliot J. Site-specific deformable imaging registration algorithm selection using patient-based simulated deformations. *Med Phys.* 2013;40(4):041911.
21. Dmitriev ID, Loo CE, Vogel WV, Pengel KE, Gilhuijs KG. Fully automated deformable registration of breast DCE-MRI and PET/CT. *Phys Med Biol.* 2013;58(4):1221–33.

22. Hardcastle N, Tome WA, Cannon DM, et al. A multi-institution evaluation of deformable image registration algorithms for automatic organ delineation in adaptive head and neck radiotherapy. *Radiat Oncol.* 2012;7:90.
23. 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(11):1901–20.
24. De Moor K, Nuyts J, Plessers L, Stroobants S, Maes F, Dupont P. Non-rigid registration with position dependent rigidity for whole body PET follow-up studies. *IEEE Nucl Sci Conf R.* 2006;6:3502–06.
25. Meijneke TR, Petit SF, Wentzler D, Hoogeman M Nuytens JJ. Reirradiation and stereotactic radiotherapy for tumors in the lung: dose summation and toxicity. *Radiother Oncol.* 2013;107(3):423–27.
26. Nithianathan S, Schafer S, Mirota DJ, et al. Extra-dimensional Demons: a method for incorporating missing tissue in deformable image registration. *Med Phys.* 2012;39(9):5718–31.
27. Al-Mayah A, Moseley J, Velec M, Hunter S, Brock K. Deformable image registration of heterogeneous human lung incorporating the bronchial tree. *Med Phys.* 2010;37(9):4560–71.
28. Lambin P, Petit SF, Aerts HJ, et al. The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother Oncol.* 2010;96(2):145–52.
29. Schreibmann E, Waller AF, Crocker I, Curran W, Fox T. Voxel clustering for quantifying PET-based treatment response assessment. *Med Phys.* 2013;40(1):012401.
30. van Elmpt W, De Ruyscher D, van der Salm A, et al. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. *Radiother Oncol.* 2012;104(1):67–71.
31. Wanet M, Lee JA, Weynand B, et al. Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: a comparison with threshold-based approaches, CT and surgical specimens. *Radiother Oncol.* 2011;98(1):117–25.
32. Bowen SR, Chappell RJ, Bentzen SM, Deveau MA, Forrest LJ, Jeraj R. Spatially resolved regression analysis of pre-treatment FDG, FLT and Cu-ATSM PET from post-treatment FDG PET: an exploratory study. *Radiother Oncol.* 2012;105(1):41–48.
33. Zips D, Zophel K, Abolmaali N, et al. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. *Radiother Oncol.* 2012;105(1):21–28.