# Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm

Catarina Veiga, Ana Mónica Lourenço, Syed Mouinuddin, Marcel van Herk, Marc Modat, Sébastien Ourselin, Gary Royle, and Jamie R. McClelland

# Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm

Catarina Veiga[a)]
*Radiation Physics Group, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom*

Ana Mónica Lourenço
*Radiation Physics Group, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom and Acoustics and Ionizing Radiation Team, National Physical Laboratory, Teddington TW11 0LW, United Kingdom*

Syed Mouinuddin
*Department of Radiotherapy, University College London Hospital, London NW1 2BU, United Kingdom*

Marcel van Herk
*Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam 1066 CX, The Netherlands*

Marc Modat and Sébastien Ourselin
*Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom*

Gary Royle
*Radiation Physics Group, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom*

Jamie R. McClelland
*Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom*

**Purpose:** The aims of this work were to evaluate the performance of several deformable image registration (DIR) algorithms implemented in our in-house software (NiftyReg) and the uncertainties inherent to using different algorithms for dose warping.

**Methods:** The authors describe a DIR based adaptive radiotherapy workflow, using CT and cone-beam CT (CBCT) imaging. The transformations that mapped the anatomy between the two time points were obtained using four different DIR approaches available in NiftyReg. These included a standard unidirectional algorithm and more sophisticated bidirectional ones that encourage or ensure inverse consistency. The forward (CT-to-CBCT) deformation vector fields (DVFs) were used to propagate the CT Hounsfield units and structures to the daily geometry for "dose of the day" calculations, while the backward (CBCT-to-CT) DVFs were used to remap the dose of the day onto the planning CT (pCT). Data from five head and neck patients were used to evaluate the performance of each implementation based on geometrical matching, physical properties of the DVFs, and similarity between warped dose distributions. Geometrical matching was verified in terms of dice similarity coefficient (DSC), distance transform, false positives, and false negatives. The physical properties of the DVFs were assessed calculating the harmonic energy, determinant of the Jacobian, and inverse consistency error of the transformations. Dose distributions were displayed on the pCT dose space and compared using dose difference (DD), distance to dose difference, and dose volume histograms.

**Results:** All the DIR algorithms gave similar results in terms of geometrical matching, with an average DSC of $0.85 \pm 0.08$, but the underlying properties of the DVFs varied in terms of smoothness and inverse consistency. When comparing the doses warped by different algorithms, we found a root mean square DD of $1.9\% \pm 0.8\%$ of the prescribed dose (pD) and that an average of $9\% \pm 4\%$ of voxels within the treated volume failed a 2%pD DD-test ($DD_{2\%\text{-pp}}$). Larger $DD_{2\%\text{-pp}}$ was found within the high dose gradient ($21\% \pm 6\%$) and regions where the CBCT quality was poorer ($28\% \pm 9\%$). The differences when estimating the mean and maximum dose delivered to organs-at-risk were up to 2.0%pD and 2.8%pD, respectively.

**Conclusions:** The authors evaluated several DIR algorithms for CT-to-CBCT registrations. In spite of all methods resulting in comparable geometrical matching, the choice of DIR implementation leads to uncertainties in dose warped, particularly in regions of high gradient and/or poor imaging quality.

## 1. INTRODUCTION

Deformable image registration (DIR) is an active and important area of research with multiple applications in adaptive radiotherapy (ART), particularly for head and neck (HN) patients.[1,2] The planning CT (pCT) can be deformed to match the daily anatomy to calculate the dose delivered per fraction,[3,4] the deformations can be applied to propagate contours,[5,6] and the fraction by fraction dose maps can be warped back to a common reference frame for summation.[7,8]

In image registration at least two input images are necessary, the source and the target images, and the result is a transformation ($T_{s\to t}$) that can be used to deform the source image ($s$) onto the target image ($t$). The majority of the research and commercial registration algorithms are unidirectional, which means they only optimize the forward transformation ($T_{s\to t}$) and do not consider the transformation in the opposite direction ($T_{t\to s}$), which can be used to deform the target image onto the source image. For applications where $T_{t\to s}$ is also required, one can simply use a unidirectional algorithm twice, by switching the roles of the source and the target images, numerically estimate the inverse of $T_{s\to t}$, or use bidirectional algorithms that optimize $T_{s\to t}$ and $T_{t\to s}$ simultaneously. In clinical applications, DIR is used to model the spatial anatomical mapping between time points; therefore, physically plausible transformations are desirable. Two concepts are associated with physically plausible deformations: inverse consistency and symmetry. Inverse-consistent registrations try to ensure that $T_{t\to s}$ is the mathematical inverse of $T_{s\to t}$ (i.e., $T_{t\to s} = T_{s\to t}^{-1}$). Symmetric registration means that identical transformations are obtained when the roles of source and target images are switched: if the source image becomes the target ($t'$, such that $t' = s$) and the target is now the source ($s'$, such that $s' = t$), a symmetric algorithm will ensure that $T_{t'\to s'} = T_{s\to t}$ and $T_{s'\to t'} = T_{t\to s}$. The two concepts are usually intertwined in the literature, but are not equal as a symmetric algorithm is not necessarily inverse-consistent (i.e., $T_{s'\to t'}$ is not guaranteed to be $T_{t'\to s'}^{-1}$), and vice-versa. The differences between symmetry and inverse consistency are more clear when considering unidirectional algorithms, since most bidirectional algorithms that aim to guarantee inverse consistency are also symmetric. For example, when performing two unidirectional registrations, one in each direction, the resulting transformations are symmetric but not inverse-consistent (switching the source and target results in the same transformations on opposite directions). If $T_{t\to s}$ is obtained by numerically estimating the inverse of the unidirectional registration result ($T_{s\to t}$), the transformations are inverse consistent but are not symmetric (estimating the inverse of the opposite transformation does not produce the same result as running the unidirectional registration in that direction).

In recent years, advanced and complex registration algorithms have been developed to be symmetric and inverse-consistent. One approach consists in using the inverse consistency error (ICE) directly as a penalty term.[9] While this encourages inverse consistency, it can reduce the ability to recover large and complex deformations, and the forward and backward transformations are only approximate inverses of each other.[10] A better, but less ad-hoc, approach is to use a stationary velocity field to parametrize the transformation.[11] This can be used to generate the deformation field in either the forward or backward direction, and these are guaranteed to be exact inverses of each other (subject to methodological approximation and numerical precision).

Dose warping and summation are important topics in ART research, since knowledge of the total dose delivered at each time point is fundamental in the ART decision-making and replanning process. Validation of DIR is a key aspect, therefore some authors have been evaluating the accuracy of DIR and dose warping using manually annotated points and structures,[5,6,12,13] physical plausibility of the transformations,[14,15] and by developing deformable and virtual phantoms with known deformations.[16–19] Others focused on estimating the accuracy requirements for dose warping[20,21] or estimating its precision.[22–25] Monte Carlo methods have also been proposed to recalculate doses using a deformed grid (with deformed and irregular voxels), which still uses DIR but minimizes the errors associated with dose interpolation.[26–28] Finally, some groups have been using dose warping to evaluate the benefits of replanning.[7,8,29] Most of the work done on these topics uses CT-to-CT registration on different anatomical sites, and/or registration algorithms that do not ensure symmetry and inverse consistency.

We investigated an ART framework for HN patients using the CT and cone-beam CT (CBCT) imaging. The Hounsfield unit (HU) information is mapped to the daily geometry for "dose of the day" calculations,[4] and the dose is remapped to the planning geometry for dose summation. Therefore, estimates of both forward and inverse transformations are required. Other groups suggest using CT and CBCT for dose remapping without requiring both transformations by calculating the dose of the day directly in the CBCT.[7,13] However, in our opinion, CBCT imaging is still unreliable for direct dose calculations, and until the image quality of CBCT is improved, a deformed CT is a good interim solution. We used four approaches to obtain those transformations, using three different DIR algorithms implemented in our in-house software (NiftyReg), all of which use a B-spline parametrization of the transformation with different

theoretically desirable properties. The aims of this work were to compare different DIR algorithms available in NiftyReg in terms of their geometrical accuracy, physical properties, and computation time and to evaluate the uncertainties inherent to using different algorithms for dose warping. Since we compare algorithms from the same software package (i.e, similar implementation), the differences found are due to the physical properties of the underlying algorithm and not due to other differences in implementation.

## 2. MATERIALS AND METHODS

### 2.A. Patient data

Retrospective data from five HN patients treated in our clinic were used in this study. The patients followed an intensity modulated radiotherapy (IMRT) treatment with a planned dose of 65 Gy delivered in 30 daily fractions. Positioning reproducibility was ensured using a personalized HN and shoulder mask with an appropriate head-rest support. The patients included in this study were selected from the few patients that were historically replanned at our institution, and therefore a challenging test cohort for image registration. They were referred for replan midtreatment when an offline review of their CBCT scans found the spinal canal or brainstem had moved out of their planning volumes, the external contour had decreased by more than 5 mm, and/or the immobilization mask no longer fitted satisfactorily.

A planning CT (GE Widebore 16 slice system) with contrast injection was acquired approximately 2 weeks before treatment and used for treatment planning. Weekly CBCTs (On-board imaging v1.4, Varian Medical Systems, Palo Alto, CA) were acquired before treatment delivery with the patient in treatment position. The CBCTs were acquired in half-fan mode, full rotation, 110 kVp, 20 mA, 20 ms, with a maximum FoV of 45 cm in diameter and 16 cm in length. Imaging resolution was $0.977 \times 0.977 \times 2.5$ mm and $0.879 \times 0.879 \times 2$ mm for the CT and CBCT scans, respectively.

### 2.B. NiftyReg

NiftyReg is an open-source registration package implemented by our institution's Centre of Medical Image Computing (http://cmic.cs.ucl.ac.uk/home/software/). It includes a block matching-based affine registration[30] and several B-spline free form deformation based algorithms.[31] NiftyReg's default algorithm is a standard unidirectional implementation on the GPU[32] and was previously used by our group for CT-to-CBCT registrations.[4] Recently more sophisticated (bidirectional, symmetric, and inverse-consistent) implementations have been incorporated in the software. An inverse-consistent symmetric[33] and a stationary velocity field transformation model implementations[34] are now freely available. NiftyReg also features a numerical estimation of the inverse of a deformation field, which uses an iterative method to estimate each vector of the inverse deformation vector field (DVF) independently using the simplex algorithm. It is similar

to other published implementations[24,35] but independently developed. This algorithm is also implemented in the GPU.

NiftyReg includes several similarity measures and regularization terms. Normalized mutual information was used for all registrations as it could successfully deal with the nonlinear relationship between CT and CBCT intensities and the spatially varying CBCT intensities. Bending energy was used as a regularization term for all of the registrations.[31] In addition, a penalty term based on the logarithm of the Jacobian determinant[36] was used for the unidirectional registrations and the inverse consistency error penalty term[33] was used for the inverse-consistent registrations. The penalty term weights were optimized for each algorithm independently. Detailed explanation of the parameters used and how they were optimized can be found in our earlier publications.[4,37,38]

### 2.C. DIR in an ART workflow

The method we propose to compute, map, and accumulate dose distributions while accounting for anatomic variations requires the pCT and CBCTs images acquired throughout the treatment. The process consists of each of the following steps (Fig. 1).

 (i) Mapping the HU information from the CT to the CBCT geometry using DIR. The process is repeated for each CBCT available for the patient.
 (ii) Dose calculations are performed on each deformed pCT (dose of the day).
(iii) Mapping the dose of the day back to the space of the pCT.
(iv) Dose distributions are accumulated and displayed on the pCT space. When a CBCT is not available for every fraction of the treatment, the weighting used when summing the doses will depend on the fractionation scheme and scans available.

The accumulated dose can potentially be used clinically to feed the ART decision-making process. At each imaging time
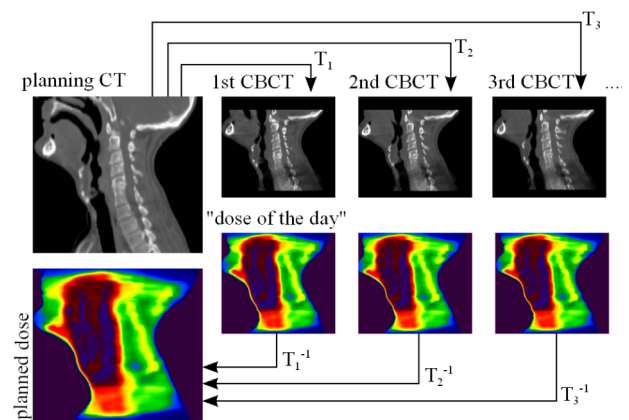


FIG. 1. Use of deformable image registration for an adaptive radiotherapy workflow. The registration maps the Hounsfield units from the CT to the daily CBCT scans, and the deformed CT is used for "dose of the day" calculations. The dose delivered is mapped back to the planning stage, where it is accumulated and the need to replan assessed.

point, the dose delivered is mapped to the same reference space, chosen to be the pCT, where it is summed to the dose from previous time points. Choosing the pCT as reference allows to iteratively compare the planned objectives with the delivered values as the treatment progresses, such that the need to replan can be regularly assessed.[7,8] The total dose delivered can also be warped to the daily geometry to assess the need to replan (using the CBCT as reference) and to feed the replanning process (using a rescan CT). In the first case, additional registrations would be required between the fractions since the reference space would change at each fraction. The latter process should be performed with utmost care as uncertainties in registration will not only affect the decision to replan but also the newly planned treatment.

In this study, we compared four different approaches to dose warping available with the NiftyReg software (we use "forward direction" to stand for the CT-to-CBCT transformations and "backward direction" to stand for the CBCT-to-CT transformations):

1. Standard asymmetric registration in both forward and backward directions ($DIR_{sas+sas}$);
2. Standard asymmetric registration in the forward direction followed by the numerical estimation of the inverse of this transformation ($DIR_{sas+inv}$);
3. Inverse-consistent symmetric registration which provides both the forward and backward transformations ($DIR_{ics}$);
4. Symmetric registration parameterized by a stationary velocity field which inherently provides both the forward and backward transformations ($DIR_{svf}$).

All the algorithms use a B-spline parametrization of the transformation with different theoretically desirable properties; while $DIR_{sas+sas}$ is symmetric but not inverse-consistent, $DIR_{sas+inv}$ is inverse-consistent but not symmetric. $DIR_{ics}$ and $DIR_{svf}$ are both symmetric, but while the first encourages inverse consistency using a penalty term, the second guarantees it by using a velocity field parametrization.[34]

## 2.D. Evaluation scheme

The different DIR algorithms were evaluated and compared in terms of geometrical matching, properties of the underlying deformations, and computation time. All the statistical tests performed on this paper were done using the MATLAB statistics toolbox and the Wilcoxon signed ranked test (95% confidence).

### 2.D.1. Geometric matching

The purpose of this part of the study was to assess the ability of the different approaches to align the same anatomical features in CT and CBCT images.

A clinical expert contoured a set of identifiable features on both the pCT and CBCT scans. The CBCT used for each patient was the last acquired before replan referral (i.e., with a relatively large deformation). The DVFs were used to propagate the features in both directions (from the pCT to the CBCT and from the CBCT to the pCT). The propagated contours were then compared with the original manually drawn contours (i.e., the features propagated from the pCT were compared with the ones manually drawn on the CBCT, and vice-versa). The features included vertebrae C1, C4, and C7 (which are subject to rigid motion only and cover the length of the cervical spinal canal), external body contour (which reflects the extent of the weight loss), and the right and left sternocleidomastoid muscles (whose deformation correlates with changes in nodal dose[39]). Common HN organs-at-risk (OARs), such as parotids and brainstem, were not considered in this evaluation because they cannot be unequivocally identified in CBCT images.

We calculated three complementary metrics that provide information about the similarity between volumes: dice similarity coefficient (DSC), false positives (FP), and false negatives (FN). DSC measures the fraction of the overlap between volumes, FP stands for the fraction of deformed volume that is not part of the manual volume, and FN stands for the fraction of manual volume that is not part of the deformed volume. Using FP and FN as well as DSC provides additional insight into the cause of geometric errors and will indicate if one structure is consistently larger/smaller than the other. To infer about the closeness between contours, the distribution of the signed Euclidean distances between the manual and deformed surfaces, also known as the distance transform (DT),[5] was calculated bidirectionally. We computed the fraction of the DT distribution larger than 2 mm ($DT_{2mm}$), mean and standard deviation of the signed DT distribution ($DT_{mean}$ and $DT_{std}$), and the 95% percentile of the DT distribution ($DT_{95\%}$). In the forward direction, DT is negative if the deformed contour is inside the manual contour, otherwise it is positive (and vice-versa for in the backward direction).

### 2.D.2. Characteristics and similarity of the deformation fields

By optimizing properly the parameters used in the different algorithms, it should be possible to obtain similar geometric matching. However, different algorithms can result in very different DVFs, particularly inside anatomy that lacks internal features or regions of increased noise and reduced contrast in the CBCT scans. Therefore, characteristics of the underlying deformations were evaluated in this section.

The smoothness of the transformations was analyzed using the harmonic energy (HE) and the properties of the determinant of the Jacobian of the transformation [det(Jac)]. The HE refers to the mean Frobenius norm of the displacement field and is inversely proportional to the smoothness of the deformation.[15] det(Jac) indicates the level of expansion/contraction at each voxel, and negative values are indicative of noninvertible and unrealistic deformations. Additionally, the ICE was calculated to investigate if the transformations were true inverses. Considering that the forward and backward transformations are inverse-consistent if their composition is equal to the identity transform, we calculated the ICE as the voxelwise difference

TABLE I. Mean values ± standard deviation of DSC, FN, FP, and DT. The results were averaged over all structures, patients, and both registration directions.

|  | $DIR_{sas+sas}$ | $DIR_{sas+inv}$ | $DIR_{ics}$ | $DIR_{svf}$ |
|---|---|---|---|---|
| DSC | $0.851 \pm 0.080$ | $0.847 \pm 0.082$ | $0.848 \pm 0.075$ | $0.851 \pm 0.073$ |
| FP | $0.17 \pm 0.11$ | $0.15 \pm 0.11$ | $0.18 \pm 0.12$ | $0.15 \pm 0.10$ |
| FN | $0.14 \pm 0.09$ | $0.15 \pm 0.10$ | $0.14 \pm 0.08$ | $0.15 \pm 0.09$ |
| $DT_{2mm}$ (%) | $9 \pm 6$ | $10 \pm 6$ | $10 \pm 6$ | $9 \pm 6$ |
| $DT_{mean}$ (mm) | $0.3 \pm 0.4$ | $0.4 \pm 0.4$ | $0.2 \pm 0.4$ | $0.3 \pm 0.4$ |
| $DT_{std}$ (mm) | $1.3 \pm 0.4$ | $1.3 \pm 0.3$ | $1.3 \pm 0.4$ | $1.3 \pm 0.4$ |
| $DT_{95\%}$ (mm) | $2.7 \pm 0.9$ | $2.7 \pm 0.9$ | $2.8 \pm 0.9$ | $2.7 \pm 1.0$ |

between the composed transformations and the identity transform.

### 2.D.3. Computation times

The time taken to complete the registrations is important for clinical translation; therefore, the computation times for each approach were measured three times per dataset. Non-GPU implementations were done on an Intel® Core™ i7-2600S CPU (2.80 GHz, 8 GB RAM), and the GPU registrations used a NVIDIA Tesla C2070 GPU card (14 multiprocessors, 6 GB dedicated memory).

### 2.D.4. Dose warping comparison

The purpose of this part of the study was to investigate the uncertainties in dose warping when using different DIR algorithms.

Dose calculations were performed on a deformed pCT and mapped back to the original pCTs using the results from the four different methods (in all cases cubic spline interpolation was used). Varian Eclipse external beam planning system analytical anisotropic algorithm was used to calculate the dose distributions using the highest available resolution (1 mm). For each patient, the same IMRT plan was applied, including beam arrangement, monitor units, and fluence maps. The dose distributions were compared within different volumes of interest using dose differences (DDs), by calculating the percentage of pixels whose DD was inferior to 2% of the prescribed dose (pD) ($DD_{2\%-pp}$), the root mean square value ($DD_{RMS}$), and the 99th percentile of the DD distribution ($DD_{99\%}$), the differences predicting the mean ($\Delta D_{mean}$) and maximum doses ($\Delta D_{max}$) to OARs, and dose volume histograms (DVHs). 2%pD was defined as tolerance criteria for DD in accordance with our internal clinical standards for the comparison of dose distributions.

The distance to dose difference (DTD) was also calculated for all the planned dose distributions (using an accuracy of 2%pD).[20] DTD is a method to estimate the required spatial accuracy of a DVF for dose warping based on the distance that one has to travel within the dose map to find a DD above a tolerance value. The DTD in different regions of interest was compared to the variability in mapping between different algorithms, by measuring voxel-by-voxel the root mean square of the DTD ($DTD_{RMS}$) and the Euclidean distance between the backward deformation vector fields (computed as a $L^2$-norm).

## 3. RESULTS

### 3.A. Geometric matching

In Table I, we present the values of DSC, FP, FN, and DT obtained averaged over both the directions of the registrations and for all patients and all structures (the ones described in Sec. 2.D.1). All the implementations provided similar results in terms of an overall geometric matching. The differences in DSC per structure were in general statistically insignificant ($p \in [0.1, 1]$) and of little clinical relevance. The two exceptions were an underperformance of $DIR_{ics}$ in the external contours ($p = 0.02$) and slightly better results for $DIR_{svf}$ in the muscles ($p < 0.01$). This is in agreement with our findings by visually inspecting the registrations. We found that $DIR_{ics}$ had difficulty in recovering larger and complex deformations, while $DIR_{svf}$ performed particularly well in the alignment of soft tissues. The effect of inverse consistency became evident when analyzing the FN and FP results: FP ≈ FN for $DIR_{sas+inv}$ and $DIR_{svf}$, as the FP in one direction coincided with the FN in the other direction, and vice-versa, and so they averaged to similar values. FP ≠ FN for $DIR_{ics}$ because even though this algorithm encourages inverse consistency, the resulting forward and backward transformations are only approximate inverses to each other.

### 3.B. Deformation field analysis

Table II shows the values of HE, det(Jac), and ICE found for each of the approaches. The voxels outside the patient were ignored when calculating the results. All the DVFs obtained had no negative det(Jac), thus were effectively invertible. In terms of physical plausibility of the DVFs, $DIR_{svf}$ provided deformations with more desirable physical properties. Lower values of HE indicate smoother transformations; therefore, the level of smoothness of the DVFs was higher in symmetric approaches ($DIR_{ics}$ and $DIR_{svf}$), which resulted in tighter intervals of the det(Jac) values. Transformations constrained with an inverse-consistent penalty ($DIR_{ics}$) considerably improved the ICE in comparison with $DIR_{sas+sas}$, but the two resulting DVFs fields were clearly not real inverses as with $DIR_{sas+inv}$ and $DIR_{svf}$, for which the mean and standard deviation of the ICE were submillimeter. Figure 2 shows the ICE map on a slice of one of the patient's anatomy. $DIR_{sas+sas}$ and $DIR_{ics}$ largest ICE values were found close to airways and in the shoulders region, where the CBCTs showed reduced contrast and higher noise. $DIR_{sas+inv}$ and $DIR_{svf}$ maximum

TABLE II. Mean values ± standard deviation for properties of the deformation vector fields: average values of HE, 1% and 99% percentiles of det(Jac), and mean, standard deviation, and 99% percentile of the ICE. The results were averaged over all patients and both registration directions.

| | $DIR_{sas+sas}$ | $DIR_{sas+inv}$ | $DIR_{ics}$ | $DIR_{svf}$ |
|---|---|---|---|---|
| HE | $0.14 \pm 0.05$ | $0.16 \pm 0.05$ | $0.11 \pm 0.05$ | $0.11 \pm 0.04$ |
| $det(Jac)_{1\%}$ | $0.53 \pm 0.12$ | $0.57 \pm 0.12$ | $0.69 \pm 0.11$ | $0.67 \pm 0.08$ |
| $det(Jac)_{99\%}$ | $1.49 \pm 0.16$ | $1.57 \pm 0.18$ | $1.33 \pm 0.13$ | $1.40 \pm 0.13$ |
| $ICE_{mean}$ (mm) | $1.6 \pm 0.7$ | $0.012 \pm 0.005$ | $0.5 \pm 0.2$ | $0.008 \pm 0.003$ |
| $ICE_{std}$ (mm) | $2.4 \pm 1.1$ | $0.04 \pm 0.03$ | $0.8 \pm 0.4$ | $0.02 \pm 0.02$ |
| $ICE_{99\%}$ (mm) | $13 \pm 6$ | $0.12 \pm 0.07$ | $4 \pm 2$ | $0.06 \pm 0.03$ |

values of ICE resulted from numerical errors when composing the DVFs in regions of high contrast boundaries.

## 3.C. Computation times

Table III shows the registration computation times measured for the different algorithms. When comparing the computation time taken to complete the registration in both directions and in the same processor, we found that $DIR_{sas+sas}$ and $DIR_{sas+inv}$ resulted in similar times, while $DIR_{ics}$ and $DIR_{svf}$ took on average 2–3 times longer. In comparison, the standard forward asymmetric registrations ran in $0.9 \pm 0.2$ min (range: 0.7–1.2 min) in the GPU. There are plans to implement all of the DIR algorithms available in NiftyReg so they can run on a GPU, but until then the use of $DIR_{ics}$ and $DIR_{svf}$ is limited to offline studies as current non-GPU computation times are too slow for online applications.

## 3.D. Dose warping comparison

Differences between the DVFs generated by different DIR algorithms will affect the final warped dose distribution. Each of the two DVF generated per patient (forward and backward) contributed to the differences in the final warped dose. First, differences in forward DVFs resulted in different deformed CTs and therefore different "doses of the day." Second, the backward DVFs remapped differently the doses of the day to the pCT space. The contribution of the first was small as we found the doses of the day to differ by less than 2% of the prescribed dose on over 95% of the body voxels.

Dosimetric differences between the results obtained with $DIR_{svf}$ and every other approach were assessed. $DIR_{svf}$ was arbitrarily chosen as the basis of this comparison since it generated the DVFs with more desirable physical properties.

Table IV presents the DD found between different methods and $DIR_{svf}$ in different regions of interest. We found no statistical evidence of any method being more similar to $DIR_{svf}$ ($p \in [0.6–1.0]$). The differences were smaller in the treated volume (TV) and larger in the 50%–95% of the prescribed dose volume (i.e., the irradiated volume that was not within the treated volume, IV-TV), where higher gradients were more likely to occur. Regions of poorer CBCT quality (low contrast and high noise within larger imaging volumes,[3] i.e., near the shoulders) within the IV ($P_{IV}$) showed higher variability between warped doses. The differences between $DD_{2\%-pp}$ between all the different identified regions were statistically significant ($p < 0.01$). Therefore, regions of higher dose gradient and poorer CBCT image quality were more prone to having larger variability in warped doses, but for different reasons. $DTD_{RMS}$ was $1.6 \pm 0.4$ mm and $2.7 \pm 0.8$ mm within IV-TV and TV, respectively. The root mean square $L^2$-norm value within IV-TV, TV, and $P_{IV}$ was $2.7 \pm 1.1$, $3.4 \pm 1.5$, and $4.2 \pm 2.1$ mm. $L^2$-norm values found for $P_{IV}$ were statistically different from other regions ($p < 0.01$), while between TV and IV-TV were not ($p = 0.1$). Voxels within IV-TV had larger DD than voxels within TV due to the local characteristics of the dose distribution (shown by DTD), while inside $P_{IV}$, the larger spatial mapping variability between DIR algorithms explained the larger DD (shown by $L^2$-norm).

The effect of the DVF in the dose mapping is complex and is theoretically expected to depend on the location of the dose gradients.[20,25] We computed the dose gradient of the planned dose distribution and related it with the values of DD within IV-TV. The correlation between gradient and DD was weak (Pearson correlation coefficient, $\rho = 0.281$), but it was clear that as the dose gradient increased the distribution of DD values became more spread and the average DD increased (Fig. 3).
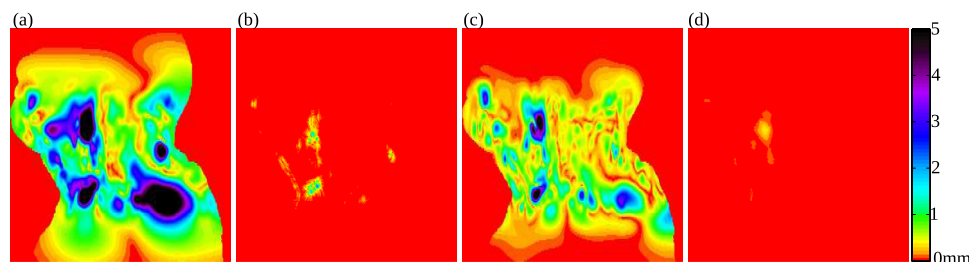


FIG. 2. ICE maps for a patient included in this study: (a) $DIR_{sas+sas}$, (b) $DIR_{sas+inv}$, (c) $DIR_{ics}$, and (d) $DIR_{svf}$.

TABLE III. Mean values ± standard deviation registration computation times (in minutes).

|  | Forward | Backward | Total |
|---|---|---|---|
| $DIR_{sas+sas}$ | 15 ± 3 | 8 ± 5 | 23 ± 6 |
| $DIR_{sas+inv}$ | 15 ± 3 | 5 ± 1 | 20 ± 3 |
| $DIR_{ics}$ | | 37 ± 12 | 37 ± 12 |
| $DIR_{svf}$ | | 61 ± 15 | 61 ± 15 |

We computed the DD, $\Delta D_{mean}$, and $\Delta D_{max}$ to the spinal canal, brainstem, and parotids (Table V). The wide range of values found for $\Delta D_{mean}$ and $\Delta D_{max}$ was indicative that larger dose differences occur depending on the particular dose distribution and relative positioning of the OARs. Analyzing the DVHs' curves obtained for each patient, we could see that, in general, all algorithms lead to similar clinical outcomes, i.e., that a replan was needed since OARs were receiving more dose than tolerated and targets less dose than planned. However, for one of the patients included in this study, while $DIR_{sas+sas}$ and $DIR_{ics}$ estimated a dose above the spinal canal tolerance, $DIR_{sas+inv}$ and $DIR_{svf}$ did not (Fig. 4). This is an example where the decision to replan could be affected by the choice of algorithm.

## 4. DISCUSSION

Regarding the geometric evaluation, all the methods resulted in good alignment between anatomical contours. In some cases, $DIR_{ics}$ had worse results than the other approaches, indicating a reduced ability to capture more complex deformations due to the introduction of additional constraint terms. The differences in geometrical alignment from $DIR_{sas+sas}$ and $DIR_{svf}$ were statistically and clinically insignificant, but the underlying properties of the deformations

TABLE IV. Mean values ± standard deviation for the dose difference test pass-percentage ($DD_{2\%-pp}$), root mean square value ($DD_{RMS}$), and the 99th percentile of the DD distribution ($DD_{99\%}$) between different approaches and $DIR_{svf}$ within different regions of interest (as a percentage of the pD). The TV corresponds to the volume encompassed by the planning 95% isodose surface, while the IV corresponds to the volume encompassed by the planning 50% isodose surface. Therefore, IV-TV is the volume where 50%–95% of the dose was planned to be delivered. $P_{IV}$ was defined as the body slices close to the shoulders (a larger imaging volume) intersected with the IV, where the CBCT Hounsfield units were not reliable.

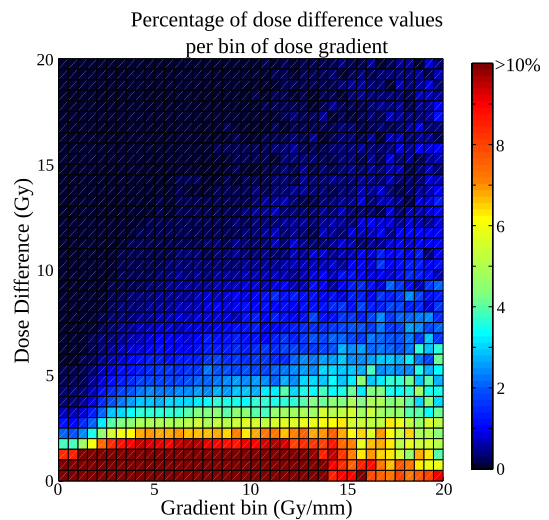|  | Method | TV | IV-TV | $P_{IV}$ |
|---|---|---|---|---|
| $DD_{2\%-pp}$ (%) | $DIR_{sas+sas}$ | 91 ± 3 | 79 ± 5 | 72 ± 7 |
| | $DIR_{sas+inv}$ | 92 ± 4 | 79 ± 6 | 70 ± 11 |
| | $DIR_{ics}$ | 91 ± 5 | 80 ± 7 | 73 ± 10 |
| $DD_{RMS}$ (%pD) | $DIR_{sas+sas}$ | 1.7 ± 0.5 | 3.9 ± 1.1 | 4.2 ± 1.3 |
| | $DIR_{sas+inv}$ | 1.6 ± 0.7 | 3.1 ± 0.8 | 4.2 ± 1.6 |
| | $DIR_{ics}$ | 1.9 ± 0.8 | 3.3 ± 0.9 | 3.8 ± 1.6 |
| $DD_{99\%}$ (%pD) | $DIR_{sas+sas}$ | 7 ± 3 | 14 ± 3 | 16 ± 4 |
| | $DIR_{sas+inv}$ | 7 ± 3 | 14 ± 4 | 19 ± 7 |
| | $DIR_{ics}$ | 9 ± 5 | 13 ± 4 | 15 ± 6 |



FIG. 3. Percentage of dose difference values per bin of dose gradient. For larger values of gradient, the DD distribution became more spread and the average DD increased.

were different. $DIR_{svf}$ resulted in deformations with more desirable physical properties, where both symmetry and inverse consistency were satisfied.

Different DIR algorithms generated different DVFs which resulted in differences when warping the dose to the planning geometry. The mean and maximum values found for DD between different DIR implementations were comparable to the values found by Salguero *et al.* when estimating the dose uncertainties of a DIR algorithm due to lack of inverse consistency.[23] In this small feasibility study, we identified situations where the choice of algorithm leads to higher uncertainties in dose warping. The first important point was the effect of the dose gradient. Where there was a high dose gradient, there was more likely a larger variability in dose between the different DIR algorithms. OARs within the high gradient region can be of concern, as we found that different methods could predict maximum doses to the spinal canal and brainstem with a difference of up to 2.8%pD. The correlation between gradient and DD was weak and similar to the values reported in the same study by Salguero *et al.* The weak correlation can be explained due to the fact that if a registration error occurs in uniform dose region the resulting dose error will be small, but when a registration error occurs in high dose gradient the resulting dose difference may be large but does not mean it will be, since there are other factors to consider besides the gradient (i.e., how large the uncertainty in spatial mapping is, whether the registration error is in the same direction as the gradient). Our findings are also in agreement with the findings of Saleh-Sayah *et al.*, as the required spatial accuracy depended on the local dose distribution.[20] A second important point is the effect of the lower quality of the CBCT in the registration uncertainty. Regions of reduced contrast and increased noise (particularly evident in larger imaging volumes like the shoulders) were more susceptible to variability in mapping between registration algorithms and therefore in larger differences between warped

TABLE V. Mean values ± standard deviation (range) for the dose difference test pass-percentage (DD$_{2\%\text{-pp}}$), root mean square value of the DD (DD$_{RMS}$), and differences predicting the mean ($\Delta D_{mean}$) and maximum doses ($\Delta D_{max}$) to the OARs.

| OAR | Method | DD$_{2\%\text{-pp}}$ (%) | DD$_{RMS}$ (%pD) | $\Delta D_{mean}$ (%pD) | $\Delta D_{max}$ (%pD) |
|---|---|---|---|---|---|
| Spinal canal | DIR$_{sas+sas}$ | 90 ± 12 (69–100) | 1.1 ± 0.6 (0.3–1.9) | 0.4 ± 0.6 (0.1–1.5) | 1.0 ± 0.9 (0.4–2.6) |
| | DIR$_{sas+inv}$ | 94 ± 5 (86–100) | 1.7 ± 1.9 (0.3–5.0) | 0.4 ± 0.6 (0.0–1.4) | 0.7 ± 0.8 (0.0–2.0) |
| | DIR$_{ics}$ | 90 ± 8 (77–96) | 1.7 ± 1.4 (0.8–4.2) | 0.5 ± 0.4 (0.2–1.1) | 0.3 ± 0.2 (0.0–0.5) |
| Brainstem | DIR$_{sas+sas}$ | 92 ± 14 (67–100) | 1.4 ± 1.9 (0.1–4.7) | 0.6 ± 0.8 (0.0–2.0) | 0.6 ± 0.4 (0.2–1.1) |
| | DIR$_{sas+inv}$ | 92 ± 11 (72–100) | 1.3 ± 1.6 (0.1–4.2) | 0.6 ± 0.8 (0.0–2.0) | 0.7 ± 1.0 (0.0–2.5) |
| | DIR$_{ics}$ | 91 ± 10 (79–100) | 1.0 ± 0.8 (0.2–2.0) | 0.5 ± 0.4 (0.1–1.1) | 1.1 ± 1.1 (0.2–2.8) |
| Parotids | DIR$_{sas+sas}$ | 86 ± 4 (79–93) | 1.5 ± 0.3 (1.1–2.0) | 0.6 ± 0.3 (0.2–1.0) | 0.2 ± 0.2 (0.0–0.5) |
| | DIR$_{sas+inv}$ | 86 ± 7 (74–95) | 1.5 ± 0.4 (1.0–2.0) | 0.3 ± 0.4 (0.0–1.4) | 0.3 ± 0.3 (0.0–1.0) |
| | DIR$_{ics}$ | 83 ± 9 (70–98) | 1.9 ± 0.8 (0.7–3.0) | 0.9 ± 0.5 (0.1–1.6) | 0.2 ± 0.2 (0.0–0.7) |

doses. The impact of the choice of DVF will depend on the dose distribution and relative positioning of OAR, and generic validation frameworks (based exclusively on geometric analysis of the deformations) are not sufficient for dose warping applications. The patients included in this study were a challenging cohort to test DIR, and it is likely that the issues reported may have less impact for less demanding patients (but then these patients will benefit less from ART). A larger study could potentially identify patterns of when different algorithms provide significant differences between doses, and flag the regions where the dose warping algorithm may be unreliable.

For an application that is sensitive to the underlying deformations such as dose warping, we believe a more complex algorithm like DIR$_{svf}$ is preferable over other approaches. This opinion is based on the similar ability to recover deformations while generating deformations with more desirable physical properties. Other studies support the theoretical advantages of ensuring symmetry and inverse consistency to improve the precision of dose warping using DIR. Bender *et al.* studied the effect of inverse consistency and transitivity in DIR for a single HN patient.[24] Lack of transitivity means that different dose distributions will be obtained depending on the order in which the registrations are used and the time

point chosen for summation. They found dose differences at OAR when different image time points were used as a reference for summation—however, when increasing the inverse consistency and transitivity those differences were considerably reduced. Inverse-consistent algorithms do not enforce transitivity making this an interesting area of research for registration developers. Yan *et al.* showed that the dose mapping inverse consistency error observed when mapping doses back-and-forth was reduced 1.5–3 times when the spatial inverse consistency was improved.[35]

The reliability of using dose warping in clinical settings is a current and open debate.[40] Our work contributes to this discussion by evaluating theoretically better DIR algorithms and investigating the uncertainties in dose warping due to the choice of algorithm in ART frameworks that use CT-to-CBCT registration. One of the main difficulties with validating dose warping is that the true point-to-point mapping is difficult or impossible to establish, especially in regions of homogeneous image intensities as is often found inside individual structures.[41,42] When new plans are based on accumulated dose, registration inaccuracies will also affect the newly planned treatment. We do not think that the difficulties inherent to validate DIR for dose warping necessarily discourage its use for clinical research, but we advise users to carefully consider
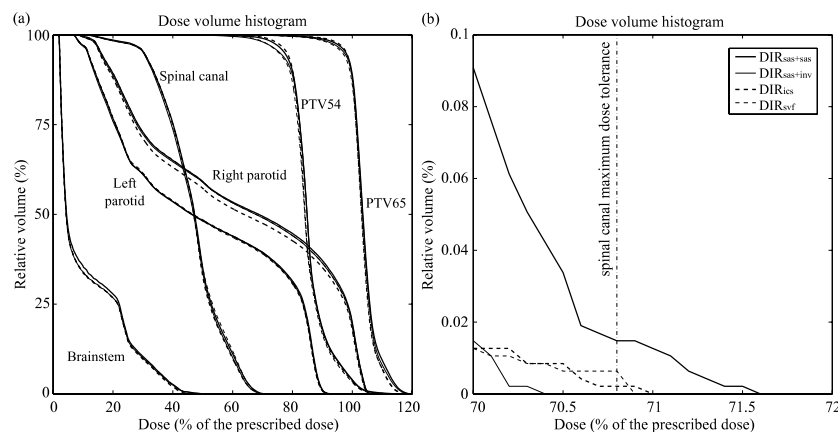


FIG. 4. Dose volume histogram obtained for a patient included in this study, using doses warped by different DIR algorithms: (a) all organs-at-risk and target volumes and (b) zoom of the maximum dose to the spinal canal. DIR$_{sas+sas}$ and DIR$_{ics}$ estimated a dose above the tolerance, while DIR$_{sas+inv}$ and DIR$_{svf}$ did not.

their choice of DIR algorithm and the conclusions that should be drawn from the results.

It should be noted that the real changes occurring in the tissue are complex and variate: sometimes tissue appears or disappears (e.g., weight loss and tumor shrinkage) and sometimes it expands/contracts or deforms in other ways. The vast majority of current DIR algorithms use a transformation model that represents expansion/contraction, but map constant image intensities, which in CT represents a constant density and so is more representative of appearing/disappearing tissue. To accurately model and recover the real physical changes that occur during a course of radiotherapy is extremely challenging but it is what an ideal DIR algorithm should be able to do and what should be an aim of the next generation of DIR algorithms. Several groups are actively working on making DIR algorithms more realistic. Examples include incorporating missing tissue in image registration by modifying existent DIR algorithms[43] and further regularizing the transformation to avoid deformation of bony anatomy.[44] However, there is still a very long way to go to achieve truly realistic DIR, and indeed, this will not just involve developing new algorithms and computational techniques but will also require a better understanding of the actual physical and biological processes that occur during a course of radiotherapy.

## 5. CONCLUSIONS

We have evaluated several DIR algorithms for CT-to-CBCT registrations and investigated the uncertainties inherent to using different DIR algorithms to warp doses to a reference geometry. Standard asymmetric and stationary vector field implementations resulted in similar geometric matching, but the properties of the DVFs were very different, with the second providing deformations with more desirable physical properties. The choice of DIR implementation had a larger impact on the dose warped in regions where the dose gradient is high and/or the CBCT image quality is poorer.

## ACKNOWLEDGMENTS

a)Author to whom correspondence should be addressed. Electronic mail: catarina.veiga.11@ucl.ac.uk

1P. Castadot, J. A. Lee, X. Geets, and V. Grégoire, "Adaptive radiotherapy of head and neck cancer," Semin. Radiat. Oncol. **20**, 84–93 (2010).

2D. L. Schwartz and L. Dong, "Adaptive radiation therapy for head and neck cancer-can an old goal evolve into a new standard?," J. Oncol. **2011**, 1–13 (2011).

3Y. Yang, E. Schreibmann, T. Li, C. Wang, and L. Xing, "Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation," Phys. Med. Biol. **52**, 685–705 (2007).

4C. Veiga, J. McClelland, S. Moinuddin, A. Lourenço, K. Ricketts, J. Annkah, M. Modat, S. Ourselin, D. D'Souza, and G. Royle, "Toward adaptive radiotherapy for head and neck patients: Feasibility study on using CT-to-CBCT deformable registration for 'dose of the day' calculations," Med. Phys. **41**, 031703 (12pp.) (2014).

5T. Zhang, Y. Chi, E. Meldolesi, and D. Yan, "Automatic delineation of on-line head-and-neck computed tomography images: Toward on-line adaptive radiotherapy," Int. J. Radiat. Oncol., Biol., Phys. **68**, 522–530 (2007).

6M. Peroni, D. Ciardo, M. F. Spadea, M. Riboldi, S. Comi, D. Alterio, G. Baroni, and R. Orecchia, "Automatic segmentation and online virtualCT in head-and-neck adaptive radiation therapy," Int. J. Radiat. Oncol., Biol., Phys. **84**, e427–e433 (2012).

7W. Lu, G. H. Olivera, Q. Chen, K. J. Ruchala, J. Haimerl, S. L. Meeks, K. M. Langen, and P. A. Kupelian, "Deformable registration of the planning image (kVCT) and the daily images (MVCT) for adaptive radiation therapy," Phys. Med. Biol. **51**, 4357–4374 (2006).

8Q. Wu, Y. Chi, P. Y. Chen, D. J. Krauss, D. Yan, and A. Martinez, "Adaptive replanning strategies accounting for shrinkage in head and neck IMRT," Int. J. Radiat. Oncol., Biol., Phys. **75**, 924–932 (2009).

9G. Christensen and H. J. Johnson, "Consistent image registration," IEEE Trans. Med. Imaging **20**, 568–582 (2001).

10B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," Med. Image Anal. **12**, 26–41 (2008).

11V. Arsigny, O. Commowick, X. Pennec, and N. Ayache, "A log-euclidean framework for statistics on diffeomorphisms," *Medical Image Computing and Computer-Assisted Intervention* (Springer, Berlin, Heidelberg, 2006), Vol. 9, pp. 924–931; see http://link.springer.com/chapter/10.1007%2F11866565_113.

12R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A. K. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," Phys. Med. Biol. **54**, 1849–1870 (2009).

13P. Risholm, J. Balter, and W. M. Wells, "Estimation of delivered dose in radiotherapy: The influence of registration uncertainty," Medical Image Computing and Computer-Assisted Intervention **14**(Pt. 1), 548–555 (2011); see http://link.springer.com/chapter/10.1007%2F978-3-642-23623-5_69.

14E. T. Bender and W. A. Tomé, "The utilization of consistency metrics for error analysis in deformable image registration," Phys. Med. Biol. **54**, 5561–5577 (2009).

15R. Varadhan, G. Karangelis, K. Krishnan, and S. Hui, "A framework for deformable image registration validation in radiotherapy clinical applications," J. Appl. Clin. Med. Phys. **14**(1), 192–213 (2013); see http://www.jacmp.org/index.php/jacmp/article/view/4066/2782.

16M. Hub, M. L. Kessler, and C. P. Karger, "A stochastic approach to estimate the uncertainty involved in b-spline image registration," IEEE Trans. Med. Imaging **28**, 1708–1716 (2009).

17A. Cherpak, M. Serban, J. Seuntjens, and J. E. Cygler, "4D dose-position verification in radiation therapy using the RADPOS system in a deformable lung phantom," Med. Phys. **38**, 179–187 (2011).

18U. J. Yeo, M. L. Taylor, J. R. Supple, R. L. Smith, L. Dunn, T. Kron, and R. D. Franich, "Is it sensible to deform dose? 3D experimental validation of dose-warping," Med. Phys. **39**, 5065–5072 (2012).

19J. Pukala, S. L. Meeks, R. J. Staton, F. J. Bova, R. R. Maon, and K. M. Langen, "A virtual phantom library for the quantification of deformable image registration uncertainties in patients with cancers of the head and neck," Med. Phys. **40**, 111703 (8pp.) (2013).

20N. K. Saleh-Sayah, E. Weiss, F. J. Salguero, and J. V. Siebers, "A distance to dose difference tool for estimating the required spatial accuracy of a displacement vector field," Med. Phys. **38**, 2318–2323 (2011).

21D. Tilly, N. Tilly, and A. Ahnesjö, "Dose mapping sensitivity to deformable registration uncertainties in fractionated radiotherapy applied to prostate proton treatments," BMC Med. Phys. **13**, 1–12 (2013).

22M. Rosu, I. J. Chetty, J. M. Balter, M. L. Kessler, D. L. McShan, and R. K. Ten Haken, "Dose reconstruction in deforming lung anatomy: Dose grid size effects and clinical implications," Med. Phys. **32**, 2487–2495 (2005).

[23]F. J. Salguero, N. K. Saleh-Sayah, C. Yan, and J. V. Siebers, "Estimation of three-dimensional intrinsic dosimetric uncertainties resulting from using deformable image registration for dose mapping," Med. Phys. **38**, 343–353 (2011).

[24]E. T. Bender, N. Hardcastle, and W. A. Tom, "On the dosimetric effect and reduction of inverse consistency and transitivity errors in deformable image registration for dose accumulation," Med. Phys. **39**, 272–280 (2012).

[25]M. J. Murphy, F. J. Salguero, J. V. Siebers, D. Staub, and C. Vaman, "A method to estimate the effect of deformable image registration uncertainties on daily dose mapping," Med. Phys. **39**, 573–580 (2012).

[26]H. Paganetti, H. Jiang, J. A. Adams, G. T. Chen, and E. Rietzel, "Monte Carlo simulations with time-dependent geometries to investigate effects of organ motion with high temporal resolution," Int. J. Radiat. Oncol., Biol., Phys. **60**, 942–950 (2004).

[27]E. Heath and J. Seuntjens, "A direct voxel tracking method for four-dimensional Monte Carlo dose calculations in deforming anatomy," Med. Phys. **33**, 434–445 (2006).

[28]H. Zhong and J. V. Siebers, "Monte Carlo dose mapping on deforming anatomy," Phys. Med. Biol. **54**, 5815–5830 (2009).

[29]B. Schaly, J. A. Kempe, G. S. Bauman, J. J. Battista, and J. Van Dyk, "Tracking the dose distribution in radiation therapy by accounting for variable anatomy," Phys. Med. Biol. **49**, 791–805 (2004).

[30]S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache, "Reconstructing a 3D structure from serial histological sections," Image Vision Comput. **19**, 25–31 (2001).

[31]D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," IEEE Trans. Med. Imaging **18**, 712–721 (1999).

[32]M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, and S. Ourselin, "Fast free-form deformation using graphics processing units," Comput. Methods Programs Biomed. **98**, 278–284 (2010).

[33]M. Modat, M. J. Cardoso, P. Daga, D. Cash, N. C. Fox, and S. Ourselin, "Inverse-consistent symmetric free form deformation," in *Biomedical Image Registration*, Lecture Notes in Computer Science Vol. 7359, edited by B. M. Dawant, G. E. Christensen, J. M. Fitzpatrick, and D. Rueckert, (Springer, Berlin, Heidelberg, 2012), pp. 79–88; see http://link.springer.com/chapter/10.1007%2F978-3-642-31340-0_9.

[34]M. Modat, P. Daga, M. Cardoso, S. Ourselin, G. Ridgway, and J. Ashburner, "Parametric non-rigid registration using a stationary velocity field," in *2012 IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA), Breckenridge, Colorado* (IEEE, 2012), pp. 145–150; see http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6164745.

[35]C. Yan, H. Zhong, M. Murphy, E. Weiss, and J. V. Siebers, "A pseudoinverse deformation vector field generator and its applications," Med. Phys. **37**, 1117–1128 (2010).

[36]M. Modat, J. R. McClelland, and S. Ourselin, "Lung registration using the NiftyReg package," in *Medical Image Analysis for the Clinic: A Grand Challenge, Workshop Proceedings from the 13th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, edited by B. van Ginneken, K. Murphy, T. Heimann, V. Pekar, and X. Den (Beijing, China, 2010), pp. 33–42.

[37]C. Veiga, J. McClelland, K. Ricketts, D. D'Souza, and G. Royle, "Deformable registrations for head and neck cancer adaptive radiotherapy," in *Image-Guidance and Multimodal Dose Planning in Radiation Therapy Workshop of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Nice, France* (2012).

[38]C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, M. Modat, S. Ourselin, D. D'Souza, and G. Royle, "Towards adaptive radiotherapy for head and neck patients: Validation of an in-house deformable registration algorithm," J. Phys.: Conf. Ser. **489**, 012083 (2014); see http://iopscience.iop.org/1742-6596/489/1/012083.

[39]V. Grégoire, P. Levendag, K. K. Ang, J. Bernier, M. Braaksma, V. Budach, C. Chao, E. Coche, J. S. Cooper, G. Cosnard, A. Eisbruch, S. El-Sayed, B. Emami, C. Grau, M. Hamoir, N. Lee, P. Maingon, K. Muller, and H. Reychler, "CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC,RTOG consensus guidelines," Radiother. Oncol. **69**, 227–236 (2003).

[40]T. E. Schultheiss, W. A. Tomé, and C. G. Orton, "It is not appropriate to deform dose along with deformable image registration in adaptive radiotherapy," Med. Phys. **39**, 6531–6533 (2012).

[41]H. Zhong, J. Kim, and I. J. Chetty, "Analysis of deformable image registration accuracy using computational modeling," Med. Phys. **37**, 970–979 (2010).

[42]U. J. Yeo, J. R. Supple, M. L. Taylor, R. Smith, T. Kron, and R. D. Franich, "Performance of 12 DIR algorithms in low-contrast regions for mass and density conserving deformation," Med. Phys. **40**, 101701 (12pp.) (2013).

[43]S. Nithiananthan, S. Schafer, D. J. Mirota, J. W. Stayman, W. Zbijewski, D. D. Reh, G. L. Gallia, and J. H. Siewerdsen, "Extra-dimensional demons: A method for incorporating missing tissue in deformable image registration," Med. Phys. **39**, 5718–5731 (2012).

[44]J. Kim, M. M. Matuszak, K. Saitou, and J. M. Balter, "Distance-preserving rigidity penalty on deformable image registration of multiple skeletal components in the neck," Med. Phys. **40**, 121907 (10pp.) (2013).