

Evaluation of deformable image registration algorithm for determination of accumulated dose for brachytherapy of cervical cancer patients

Reza Mohammadi, MSc^{1,2}, Seied Rabi Mahdavi, PhD¹, Ramin Jaber, PhD³, Zahra Siavashpour, PhD⁴, Leila Janani, PhD⁵, Ali Soleimani Meigooni, PhD⁶, Reza Reiazi, PhD^{1,2}

¹Department of Medical Physics, Iran University of Medical Sciences, Tehran, Iran, ²Medical Image and Signal Processing Research Core, Iran University of Medical Sciences, Tehran, Iran, ³Radiation Oncology Research Center, Tehran University of Medical Sciences, Tehran, Iran,

⁴Department of Medical Radiation Engineering, Shahid Beheshti University, Tehran, Iran, ⁵Department of Biostatistics, Iran University of Medical Sciences, Tehran, Iran, ⁶Comprehensive Cancer Centers of Nevada, Las Vegas, NV, United States

Abstract

Purpose: This study was designed to assess the dose accumulation (DA) of bladder and rectum between brachytherapy fractions using hybrid-based deformable image registration (DIR) and compare it with the simple summation (SS) approach of GEC-ESTRO in cervical cancer patients.

Material and methods: Patients ($n = 137$) with cervical cancer treated with 3D conformal radiotherapy and three fractions of high-dose-rate brachytherapy were selected. CT images were acquired to delineate organs at risk and targets according to GEC-ESTRO recommendations. In order to determine the DA for the bladder and rectum, hybrid-based DIR was done for three different fractions of brachytherapy and the results were compared with the standard GEC-ESTRO method. Also, we performed a phantom study to calculate the uncertainty of the hybrid-based DIR algorithm for contour matching and dose mapping.

Results: The mean \pm standard deviation (SD) of the Dice similarity coefficient (DICE), Jaccard, Hausdorff distance (HD) and mean distance to agreement (MDA) in the DIR process were 0.94 ± 0.02 , 0.89 ± 0.03 , 8.44 ± 3.56 and 0.72 ± 0.22 for bladder and 0.89 ± 0.05 , 0.80 ± 0.07 , 15.46 ± 10.14 and 1.19 ± 0.59 for rectum, respectively. The median (Q1, Q3; maximum) Gy_{EQD2} differences of total D_{2cc} between DIR-based and SS methods for the bladder and rectum were reduced by -1.53 (-0.86 , -2.98 ; -9.17) and -1.38 (-0.80 , -2.14 ; -7.11), respectively. The mean \pm SD of DICE, Jaccard, HD, and MDA for contour matching were 0.98 ± 0.008 , 0.97 ± 0.01 , 2.00 ± 0.70 and 0.20 ± 0.04 , respectively for large deformation. Maximum uncertainty of dose mapping was about 3.58%.

Conclusions: The hybrid-based DIR algorithm demonstrated low registration uncertainty for both contour matching and dose mapping. The DA difference between DIR-based and SS approaches was statistically significant for both bladder and rectum and hybrid-based DIR showed potential to assess DA between brachytherapy fractions.

J Contemp Brachytherapy 2019; 11, 5: 469–478

DOI: <https://doi.org/10.5114/jcb.2019.88762>

Key words: deformable image registration, brachytherapy, cervix, HDR.

Purpose

Combining external beam radiotherapy (EBRT) and chemotherapy with high-dose-rate brachytherapy (HDR-BT) as boost is the standard approach to treat locally advanced cervical cancer. For HDR-BT, several types of intracavitary applicators are commonly used with or without interstitial needles. After completing the applicator insertion in each fraction, image-guided treatment planning is used to calculate three-dimensional (3D) dose distribution. Some recommendations have been published by the Groupe Européen de Curiethérapie-European Society for Therapeutic Radiology and On-

cology (GEC-ESTRO) working group, and the American Brachytherapy Society (ABS) based on 3D image-guided treatment planning for contouring, applicator reconstruction, prescribing, and reporting [1,2,3,4,5,6]. The recommended prescription dose for high-risk clinical target volume (HR-CTV) with $\alpha/\beta = 10$ is 80-90 Gy_{EQD2} (45-50 Gy in 1.8-2.0 Gy/fraction as EBRT), expressed as the equivalent dose in 2 Gy per fraction (EQD_2). The maximum dose to 2 cm^3 (D_{2cc}) of the organs at risk (OARs) such as bladder, rectum and sigmoid with $\alpha/\beta = 3$ should not exceed 90 Gy_{EQD2} , 75 Gy_{EQD2} , and 75 Gy_{EQD2} , respectively.

The GEC-ESTRO recommended simple summation (SS) of EBRT dose prescription and D_{2cc} to predict toxicity

Address for correspondence: Reza Reiazi, PhD, Department of Medical Physics, Iran University of Medical Sciences, Shahid Hemmat Highway, Tehran, 1449614535, Iran, phone: +982186704578, fax: +982188622647, Accepted: 17.09.2019
 ✉ e-mail: reiazi.r@iums.ac.ir Published: 30.10.2019

in OARs. This is based on the assumption (the worst-case assumption) that each volume of interest has the same location in EBRT and during each HDR-BT fraction. Some studies have indicated that organ motions, differences in filling of bladder/rectum, presence of gas in the rectum, and differences in location of applicator cause the SS approach to overestimate the total dose in OARs, thereby limiting elevation of the HR-CTV dose or target coverage [7,8,9,10].

Recently, some investigations have proposed a new strategy based on image registration to improve calculation of maximum dose to OARs. In several studies, rigid image registration (RIR) has been applied to assess dose accumulation (DA); however, the complexity of organs' deformation due to the anatomic changes and applicator insertion limits the accuracy of calculation [11,12,13]. To overcome these limitations, several deformable image registration (DIR) algorithms have been investigated [14,15,16,17,18]. Validation of the accuracy of DIR algorithms is a challenging task due to complexity, clarity issues, and uncertainties in the image. There are some methods to evaluate the accuracy of DIR including propagation of the same points or contours between registered images, registration of images of physical or virtual phantoms, and patient image registration [19].

On the other hand, high accuracy of the DIR algorithm does not necessarily address the accuracy of DA. Hayashi *et al.* [20] estimated the DIR accuracy by creating a virtual phantom and using the Dice similarity coefficient (DICE). Although they reported DICE equal to 0.96 in the virtual phantom, the average dose difference was 8%, which was associated with uncertainty of the DIR algorithm. Reniers *et al.* [21] also reported 5-10% differences between DA using DIR and GEC-ESTRO recommendations depending on the DIR scheme. However, their study did not examine a sufficient number of patients.

To the best of our knowledge, there is no study investigating the effect of the DIR algorithm in terms of dosimetry by virtual phantom and a subsequent statistical population of patient data to obtain reasonable results for comparing DIR and SS approaches for DA. The aim of this study is to assess DA of bladder and rectum using a hybrid-based DIR algorithm and compare it statistically with the GEC-ESTRO method in multiple fractions of cervical cancer brachytherapy (BT).

Material and methods

This study was conducted based on two main parts including virtual phantom and patient data.

Virtual phantom data

To investigate the uncertainty of the hybrid-based DIR algorithm, we created a virtual phantom as ground truth data by commercial software ImSimQA v4.0 (Oncology Systems Limited, UK) as displayed in Figure 1A-C. "ImSimQA tool can serve as a virtual deformable quality assurance tool by simulating clinically observed organ deformations" [22].

In this study, we simulated a simple human pelvis. In order to produce warped image sets based on applying

geometric deformation, there are two approaches: global and local deformation. Both global and local deformation are based on radial basis function (RBF) and different kernel functions.

Global deformations

The ImSimQA uses RBF with thin-plate spline (TPS) as an interpolation to create global deformation. In this regard, two sets of points must be chosen manually on the images including source points (SPs) and target points (TPs). This tool creates a grid in each slice of image and then assigns x , y to all points (SPs and TPs) which are manually selected by clicking with the mouse on the image. Vectors with origin from SPs to TPs determine deformation direction of the grid. Then mapping function $f(x, y)$ which defined RBF with TPS kernels in the software will map SPs to TPs to create global warped image sets. More details about the $f(x, y)$ function can be found in the Bookstein *et al.* study [23].

Local deformations

ImSimQA uses RBF with compact support radial basis function (CSRBF) as an interpolation to create local deformation. This will be applied on the image same as the global deformation process with the only difference in kernels. More details about CSRBF can be found in the study by Wendland *et al.* [24].

We created two groups of global and local warped image sets. We considered average 5 mm and 15 mm distances between SPs and TPs to create groups of warped image sets named small and large warped image sets, respectively (Figure 1E-H).

Threshold-base segmentation was performed on image sets prior to registration by the commercial software Artiview v3.20.1 (Aquilab, France) as illustrated in Figure 1D-H. All images were imported to the SagiPlan (Eckert & Ziegler BEBIG Co., Germany) BT treatment planning system. A tandem and ovoid applicator (Fletcher type) with the same position, size, length, dwell position, and dwell time was added to create the same isodose curves across all virtual phantoms.

We divided the designed phantoms to primary and secondary categories for the registration process. The images of the phantom without deformation were considered as primary (fixed or reference images), whilst images of the warped phantom were regarded as secondary images (moving or target images). RIR followed by DIR was performed between the primary and each secondary image to calculate the contour matching and dose mapping. When performing DIR between two image sets, a DIR algorithm without uncertainty should generate a new image set with the same structures (contour matching) and dose distribution (dose mapping) equal to the primary image.

Patient data

Data for $n = 137$ patients with locally advanced cervical cancer treated with 3D conformal radiotherapy (3DCRT) and HDR-BT from 2016 to 2018 were included in this study. All patients were treated with 3DCRT with

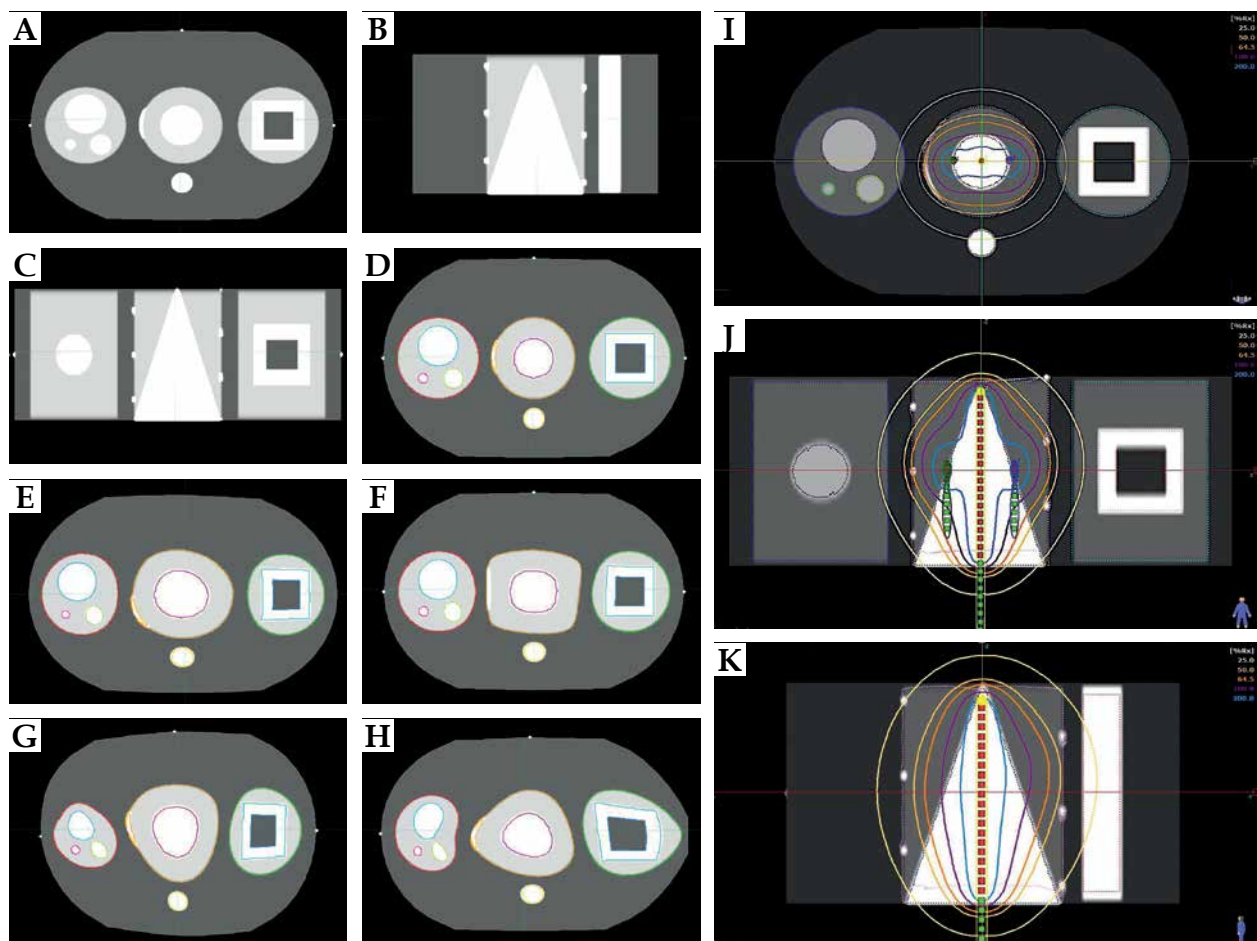


Fig. 1. One representative slice of created virtual phantom: (A) axial view, (B) sagittal view, (C) coronal view. The deformed phantoms are presented in (E-H). (D) shows virtual phantom without deformation with defined structures by threshold-based segmentation. (E) and (F) are small warped images. (G) and (H) are large warped images. Both categories, small and large warped image sets, include global and local deformation. (E) and (G) are global warped images. (F) and (H) are local warped images. Both global and local warped image sets were created by RBF function with TPS and CSRBF kernel function, respectively. (I), (J) and (K) show virtual phantom without deformation with applicator in place and dose distribution in different view. 100 percentage of dose is related to 8.6 Gy

50 Gy in 25 fractions and 3 fractions of HDR-BT (8.6 Gy/Fr) [25] using tandem and ovoids (Fletcher type without interstitial needles). The 3DCRT was planned to cover the clinical target volume (CTV) comprising the cervix, entire uterus, bilateral parametria, upper half of vagina, and lymph nodes (common, external and internal iliac) by at least 95% of the prescription dose. Before each fraction of BT, a Foley catheter filled with contrast agent (7cc) was inserted into the bladder. 120 cm³ of normal saline solution was injected into the bladder and clamped immediately before the CT scanning. Thereafter, CT images (GE HiSpeed Dual-Slice) with 3 mm slice thickness were acquired to delineate OARs and the target to calculate the dose distribution according to GEC-ESTRO recommendations. Also, MRI images (T2 weighted) were acquired to determine HR-CTV precisely just before the first BT fraction. HDR-BT was planned using SagiPlan as the treatment planning system and a ⁶⁰Co source (Eckert & Ziegler BEBIG Co., Germany) to deliver 90% of the dose to the HR-CTV. We considered dose constraints for

OARs including bladder, rectum, and sigmoid according to GEC-ESTRO recommendations.

Deformable image registration and dose accumulation

In this study, we applied a hybrid-based DIR algorithm with sum of the squared differences (SSD) as the similarity measure by a commercial system (MIM v6.7; MIM software, Inc.). The hybrid-based DIR algorithm of this software considered both intensity and contours (should be selected manually) of images. The cost function term contains two parts: $C_{similarity}$ and C_{smooth} .

$$cost\ function = C_{similarity} + \lambda C_{smooth}$$

$C_{similarity}$ minimizes the SSD between primary and secondary images. C_{smooth} is a regularization term that maintains smoothness of the control point grid and the resulting DVF. λ is the balance of these two terms, which by default has been set to a value of 0.5.



Fig. 2. OARs contours masked to value 1000, -1000 and 0 for bladder, rectum and sigmoid, respectively

Initially, we performed rigid registration to align images (using the whole image), after which DIR was performed. This approach was done for both virtual phantom and patient images. Contours of bladder, rectum, and sigmoid from patient data and all structures in the virtual phantom were selected for the DIR process. Also, RIR and DIR process durations were calculated.

In order to remove the rectal gas and bladder Foley catheter filling effects in the DIR process, we masked contours to the values 1000, -1000, and 0 for bladder, rectum, and sigmoid, respectively (Figure 2).

Patient's CT image data of the first fraction (Fr1) were considered as primary, whilst second (Fr2) and third fraction (Fr3) images were considered as secondary images in the registration process. Displacement vector fields obtained from the DIR process (from warped Fr2 to Fr1 and warped Fr3 to Fr1) were applied to the dose distribution for DA. All doses were converted to EQD₂ using a linear-quadratic model with $\alpha/\beta = 3$ for normal tissues (OARs). DVH parameters including $D_{0.1cc}$, D_{1cc} , D_{2cc} and D_{5cc} of bladder and rectum were calculated by accumulated DVH and SS approaches. Also, in the virtual phantom, $D_{0.1cc}$, D_{1cc} , D_{2cc} and D_{5cc} between structures of two registered images were used to calculate the dose mapping.

The overall framework of this study is summarized in Figure 3.

Deformable image registration accuracy

Different metrics have been developed to evaluate of DIR accuracy [26,27,28]. We used DICE, Jaccard, Hausdorff distance (HD), and mean distance to agreement (MDA) metrics between structures to quantitatively calculate the uncertainty of contour matching in the virtual phantom. We also calculated these metrics for patient data. DICE and Jaccard are given as:

$$DICE = 2 \frac{|A \cap B|}{|A| + |B|}$$

$$Jaccard = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

DICE and Jaccard measure the overlap of two contours where A and B denote the contours of primary and secondary images, respectively. DICE and Jaccard range from 0 to 1. If A and B completely overlapped, DICE and

Jaccard would be 1; otherwise it would be 0 for complete non-overlapping conditions.

HD describes the maximum closest distance for each apex of the two volumes (mm). Also, MDA is defined as the average distance of each point on one contour to the closest points on the other (mm). These two parameters can also be measured as below:

$$HD(X, Y) = \sup_{y \in Y} \inf_{x \in X} d(X, Y)$$

$$MDA(X, Y) = \text{avg} \inf_{y \in Y} d(X, Y)$$

Where sup, inf, and avg stand for supremum, infimum, and average [29]. These metrics were calculated for both RIR and DIR processes.

Statistical analysis

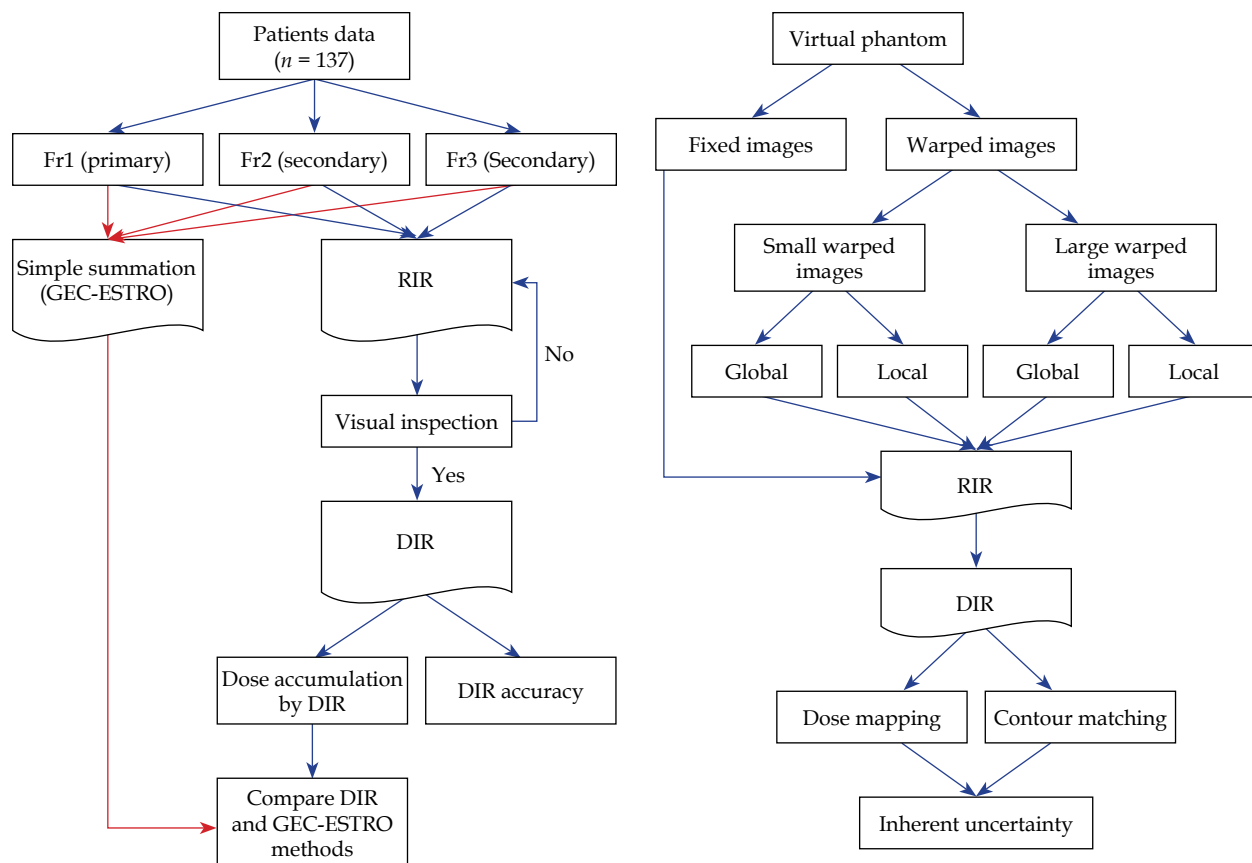
DVH parameters from DIR-based and SS approaches were compared using the Wilcoxon signed rank sum test. IBM SPSS v23 (SPSS, Inc.) was used for all statistical analyses, while mean \pm standard deviation (SD) or the median (Q_1 , Q_3 ; maximum) was used for summarizing the data. Interquartile range (IQR) was also employed to find outliers in the data, which equaled differences between the third and first quartiles. Three criteria were considered including 1.5 IQR as a mild outlier, 2 IQR and 2.5 IQR as moderate outliers, and 3 IQR as an extreme outlier. A p -value less than 0.05 was considered statistically significant.

Results

Virtual phantom

Table 1 presents the uncertainty of contour matching and dose mapping in the DIR process. The results are shown for all metrics (DICE, Jaccard, HD, MDA). DIR has the highest accuracy for contour matching, where the differences between DIR and RIR were significant ($p < 0.001$ for all metrics).

For, $D_{0.1cc}$, D_{1cc} , D_{2cc} and D_{5cc} between two registered images (primary and each secondary image), the overall mean \pm SD difference of the dose distribution among all structures in the DIR process was 0.012 ± 0.008 Gy and



Fr – fraction, RIR – rigid image registration, DIR – deformable image registration

Fig. 3. Flowchart of entire procedure used in this study. Virtual phantom was created as ground truth data to calculate uncertainty of DIR algorithm. RIR followed by DIR was performed between fixed image and each warped image set. When performing DIR between two image sets, a DIR algorithm without uncertainty should generate a new image set with the same structures (contour matching) and dose distribution (dose mapping) equal to the primary image. Patient’s CT image data sets for first fraction (Fr1) were considered as primary whilst second (Fr2) and third fraction (Fr3) images were considered as secondary images in the registration process, respectively. RIR was applied to align images, then DIR was performed. Displacement vector fields obtained from DIR process were applied to dose distribution for dose accumulation. All doses were converted to EQD₂ using linear-quadratic model with “ $\alpha/\beta = 3$ ” for normal tissues. DVH parameters such as D_{0.1cc}, D_{1cc}, D_{2cc} and D_{5cc} of bladder and rectum from accumulated DVH and simple summation approaches were calculated and results were compared

0.043 ± 0.013 Gy, leading to 1.05 ± 0.26% and 3.58 ± 0.83% for small and large deformations, respectively. These results indicated that the maximum uncertainty of D_{2cc} for the DIR method was about 3.58%.

Patients

The mean duration of RIR and DIR processes was 20 and 50 seconds, respectively. The mean ± SD values of DICE, Jaccard, HD, and MDA in RIR and DIR processes are reported in Table 2. There were statistically significant differences between the metrics calculated by DIR compared to RIR ($p < 0.001$ for all metrics), where DIR had the highest accuracy to match contours for bladder and rectum.

The median (Q₁, Q₃; maximum) Gy_{EQD2} of total D_{2cc} for the bladder and rectum by the SS approach was 38.64 (31.93, 44.51; 49.58) and 18.66 (15.73, 21.26; 30.56), respectively. After DIR was performed, the median (Q₁, Q₃; maximum) Gy_{EQD2} of total D_{2cc} for the bladder and rec-

tum was 36.38 (30.42, 40.17; 45.96) and 17.11 (14.48, 19.51; 27.36), respectively. The median (Q₁, Q₃; maximum) Gy_{EQD2} differences of total D_{2cc} between DIR-based and SS approaches for the bladder and rectum were reduced to -1.53 (-0.86, -2.98; -9.17) and -1.38 (-0.80, -2.14; -7.11), respectively (Figure 4).

The differences of D_{0.1cc}, D_{1cc}, D_{2cc}, and D_{5cc} between the DIR and SS approaches for bladder and rectum are outlined in Table 3. The DIR-based DA was lower than SS’s DA for all patients (except for D_{0.1cc} in some patients). The differences of D_{0.1cc}, D_{1cc}, D_{2cc}, and D_{5cc} between DIR and SS were found to be significant for both bladder ($p < 0.001$) and rectum ($p < 0.001$). The outlier data for differences of D_{2cc} between DIR-based and SS methods are summarized in Table 4. Details of the results for D_{0.1cc}, D_{1cc}, and D_{5cc} can be found in the supplements. The volume differences for patients who were moderate outlier data of bladder and rectum were 90 and 53 cc, respectively (Table 4). The median volume differences for bladder

Table 1. Uncertainty of contour matching and dose mapping in DIR process

	Small deformation		Overall	Large deformation		Overall
	Global	Local	Mean \pm SD	Global	Local	Mean \pm SD
DICE _{RIR}	0.71 \pm 0.25	0.97 \pm 0.02	0.84 \pm 0.22	0.71 \pm 0.08	0.76 \pm 0.21	0.74 \pm 0.16
Jaccard _{RIR}	0.60 \pm 0.24	0.95 \pm 0.04	0.77 \pm 0.24	0.56 \pm 0.10	0.66 \pm 0.21	0.61 \pm 0.17
HD _{RIR} (mm)	9.53 \pm 3.59	2.66 \pm 1.53	6.09 \pm 4.40	14.59 \pm 4.57	13.18 \pm 5.00	13.89 \pm 5.05
MDA _{RIR} (mm)	3.58 \pm 1.69	0.30 \pm 0.26	1.94 \pm 2.03	4.34 \pm 1.08	2.87 \pm 1.06	3.60 \pm 1.30
DICE _{DIR}	0.99 \pm 0.007	0.99 \pm 0.008	0.99 \pm 0.008	0.98 \pm 0.006	0.98 \pm 0.010	0.98 \pm 0.008
Jaccard _{DIR}	0.97 \pm 0.01	0.98 \pm 0.01	0.97 \pm 0.15	0.97 \pm 0.013	0.97 \pm 0.02	0.97 \pm 0.01
HD _{DIR} (mm)	1.61 \pm 0.71	1.67 \pm 0.88	1.64 \pm 0.80	1.95 \pm 1.00	2.00 \pm 0.79	2.00 \pm 0.70
MDA _{DIR} (mm)	0.18 \pm 0.04	0.15 \pm 0.04	0.16 \pm 0.04	0.20 \pm 0.04	0.20 \pm 0.03	0.20 \pm 0.04
D _{0.1cc} (Gy)	0.012 \pm 0.004	0.010 \pm 0.001		0.058 \pm 0.006	0.041 \pm 0.006	
D _{1cc} (Gy)	0.020 \pm 0.013	0.009 \pm 0.006	0.012 \pm 0.008	0.052 \pm 0.012	0.037 \pm 0.010	0.043 \pm 0.013
D _{2cc} (Gy)	0.014 \pm 0.007	0.010 \pm 0.005		0.047 \pm 0.012	0.031 \pm 0.014	
D _{5cc} (Gy)	0.015 \pm 0.007	0.009 \pm 0.006		0.052 \pm 0.006	0.036 \pm 0.006	

RIR – rigid image registration, DIR – deformable image registration, DICE – dice similarity coefficient, HD – Hausdorff distance, MDA – mean distance to agreement. Differences between RIR and DIR metrics were significant ($p < 0.001$)

Table 2. DIR accuracy metrics for bladder and rectum

	Bladder	Rectum
	Mean \pm SD	
DICE _{RIR}	0.72 \pm 0.11	0.60 \pm 0.11
Jaccard _{RIR}	0.58 \pm 0.12	0.43 \pm 0.11
HD _{RIR} (mm)	18.96 \pm 7.58	24.45 \pm 10.66
MDA _{RIR} (mm)	4.28 \pm 2.20	5.11 \pm 2.39
DICE _{DIR}	0.94 \pm 0.02	0.89 \pm 0.05
Jaccard _{DIR}	0.89 \pm 0.03	0.80 \pm 0.07
HD _{DIR} (mm)	8.44 \pm 3.56	15.46 \pm 10.14
MDA _{DIR} (mm)	0.72 \pm 0.22	1.19 \pm 0.59

RIR – rigid image registration, DIR – deformable image registration, DICE – Dice similarity coefficient, HD – Hausdorff distance, MDA – mean distance to agreement

and rectum for extreme outlier data were 103 and 79 cc, respectively. The median volume differences for bladder and rectum for all patients were 50 and 20 cc, respectively.

Discussion

This study is the first to use hybrid-based DIR with a virtual phantom and subsequent statistical population of patient data to evaluate uncertainty and DA, respectively for bladder and rectum in cervical cancer patients. We observed that hybrid-based DIR had 0.16 mm and 0.2 mm uncertainty for small and large deformations, respectively. This resulted in 1.05% and 3.58% uncertainty in dose mapping when the contours overlapped. In com-

parison with Hayashi *et al.* [20], our hybrid-based DIR algorithm was found to have a lower uncertainty (3.58% vs. 8% reported by Hayashi *et al.*). Concerning patient data, we found that the differences of DIR-based and SS approaches were statistically significant for D_{0.1cc}, D_{1cc}, D_{2cc} and D_{5cc} for both bladder ($p < 0.001$) and rectum ($p < 0.001$), indicating that the hybrid-based DIR algorithm has a potential for DA between brachytherapy fractions.

There are two general categories to evaluate uncertainty of the DIR algorithm: physical and virtual phantoms. The virtual phantom is the preferred method, since the physical phantom suffers from some limitations including inability to know the transformation of each voxel as well as difficult and time-consuming creation of each clinical situation [19]. In addition, different noise patterns can be added to virtual phantom images to simulate noise variation between the image sets [19]. It is possible to apply a known geometric function to the virtual phantom for creating ground truth data. Nevertheless, it should be noted that applying a geometric deformation function and DIR algorithm should not be the same model as it causes bias in the results. In our study, we applied RBF with different kernels as a geometric deformation function (RBF is landmark-based transformation function) to the virtual phantom to create local and global warped image sets. We used hybrid-based DIR to resolve the deformation, which is a different model from RBF with different kernels.

It is important to choose suitable DIR algorithms. Some investigations have applied RIR and intensity-based DIR for DA [11, 13]. Due to a lack of degree of freedom for the rigid algorithm, it is impossible to perform local mapping. Sabater *et al.* [30] investigated DA by RIR for

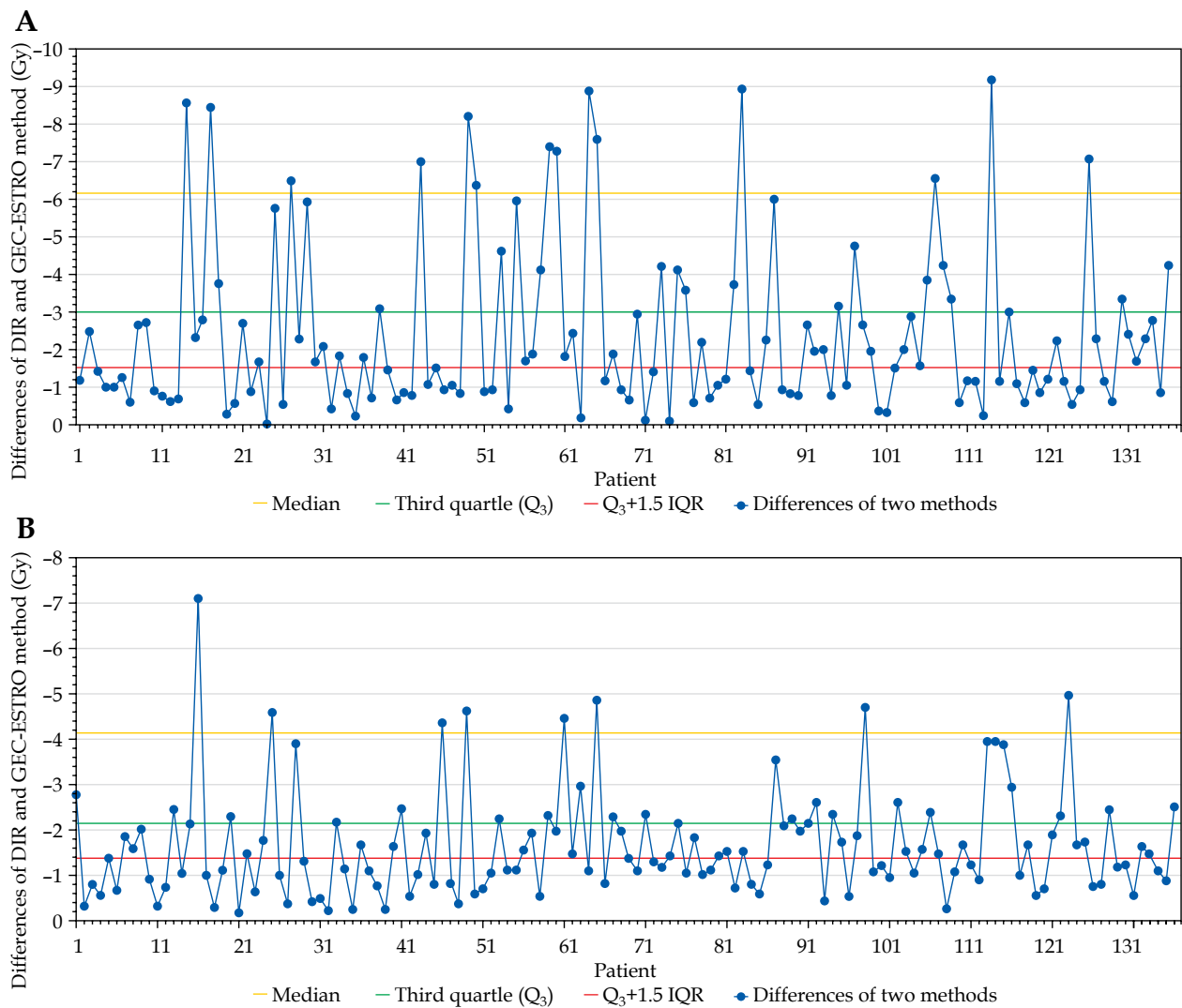


Fig. 4. Differences of DA for D_{2cc} between DIR-based and SS methods of **A)** bladder and **B)** rectum for individual patients

patients treated with 3 or 5 BT fractions. They found that there were no significant differences in terms of $D_{0.1cc}$, D_{1cc} , D_{2cc} and D_{5cc} between SS and RIR-based summation for bladder and rectum in both BT fractionation regimes. Greater differences were found with 5 versus 3 fractions for RIR, suggesting that with an increased number of treatment fractions, the uncertainty due to interfractional changes grows. Jamema *et al.* [31] compared the intensity and contour-based DIR for DA. They found that the contour-based approach had greater accuracy compared to the intensity-based DIR. The intensity-based algorithm does not include a regularization term to consider the mechanical properties of OAR walls, causing systematic underestimation of the dose. In addition, the presence of a Foley catheter in the bladder and gas in the rectum affect the DIR process, which can be eliminated by masking contours to the special Hounsfield unit number.

Validation of image registration is a challenging task. There are several methods to validate the image registration including propagation of the same points or contours between registered images, registration of images

of physical or virtual phantoms, and patient image registration. Each of them involves certain issues including construction of a physical phantom and design of a virtual phantom to simulate complexity of the real situation, uncertainty in identifying boundary of contours, finding fiducial markers accurately, effect of sample points,

Table 3. Differences of DA between DIR-based and SS methods for bladder and rectum

	Bladder	Rectum
	Median (Q_1, Q_3)	
$D_{0.1cc}$ (Gy)	-4.08 (-1.59, -9.69)	-2.26 (-0.97, -4.64)
D_{1cc} (Gy)	-2.01 (-1.07, -4.15)	-1.42 (-0.74, -2.56)
D_{2cc} (Gy)	-1.53 (-0.86, -2.98)	-1.38 (-0.80, -2.14)
D_{5cc} (Gy)	-1.18 (-0.65, -2.63)	-1.11 (-0.71, -1.74)

Q_1 – first quartile, Q_3 – third quartile
Differences of DA between DIR-based and SS approaches were significant ($p < 0.001$)

Table 4. Outlier data for differences of D_{2cc} between DIR-based and SS methods

	Bladder	Number of included (outlier) patients	Percentage (%)	Rectum	Number of included (outlier) patients	Percentage (%)
IQR	-2.12 (Gy)			-1.33 (Gy)		
$Q_3 + 1.5$ IQR	-6.16 (Gy)	123 (14)	89.8 (10.2)	-4.125 (Gy)	129 (8)	94.2 (5.8)
$Q_3 + 2$ IQR	-7.22 (Gy)	128 (9)	93.4 (6.6)	-4.79 (Gy)	134 (3)	97.8 (2.2)
$Q_3 + 2.5$ IQR	-8.28 (Gy)	132 (5)	96.4 (3.6)	-5.455 (Gy)	136 (1)	99.3 (0.7)
$Q_3 + 3$ IQR	-9.34 (Gy)	0	0	-6.12 (Gy)	136 (1)	99.3 (0.7)

IQR – interquartile range, Q_3 – third quartile, total patients – 137

effect of voxel size and slice thickness and uncertainties of imaging devices which is caused by lack of a robust and reliable validation method especially for DA. In the absence of ground truth, robustness and consistency tests provide quantitative validation of the image registration. Robustness and precision in the registration outcomes is established by testing the bias and sensitivity of the results after adding random noise or choosing different a priori or parameter settings. Consistency evaluates the performance of a registration algorithm in circular transformations. The deformation vector field in DIR can be tested using an inverse consistency test between image A and B, where A is deformed to match B and B is separately matched to A. An inverse-consistent algorithm will produce a true inverse deformation field when the moving and fixed images are switched.

In a study by Reniers *et al.*, the average distance between contours of 2 mm or less was suggested for BT dosimetry [21]. In this study, we obtained 0.72 mm and 1.19 mm for bladder and rectum, respectively, which are better results compared to Reniers *et al.*'s suggestion. Several previous studies did not find any significant differences between DIR and SS approaches, which might be related to the patient sample size, number of treatment fractions, or DIR algorithm used [21,30,32]. We considered values above $Q_3 + 1.5$ IQR as outlier data (Table 4) and found statistically significant differences at least for 90% of patients between DIR-based and SS approaches due to a reasonable sample size and hybrid-based DIR algorithm with high accuracy.

There are many uncertainties in BT including organ motions (inter- or intra-fractions), applicator displacement after CT scan, differences in filling of bladder/rectum, presence of gas in the rectum, and uncertainty in dose calculation using TG-43 formalism, affecting evaluation of the delivered dose [33,34]. Changes in the bladder filling and gas in rectum between fractions cause displacements of 0.1 cc, 1 cc, 2 cc, and 5 cc volumes of OARs which deliver high doses [15,35]. One patient ($n = 114$) had a greater volume difference (154 cc) in the bladder compared to the first fraction of bladder volume. DIR-based DA was -9.18 Gy less than SS, which may have occurred due to large inter-fraction volume differences. In other words, this case had a small bladder volume in the first fraction and a large bladder volume in the second and third fractions. Thus, the DIR process shrinks hotspot

points from large contour to small contour, thereby reducing the D_{2cc} value. Andersen *et al.* [32] observed that volume differences and D_{2cc} were significant for bladder. Bladder volume causes changes in the maximum dose to the rectum and sigmoid, and the interfraction volume difference is important in the DIR process. Siavashpour *et al.* [35] suggested an optimum volume for the bladder (70 cc) which decreases doses to OARs. We also suggest 70 cc for the bladder as it reduces the dose to OARs followed by a diminished interfraction volume difference, improving DIR accuracy as reported by a previous study [15]. The DIR-based DA of the rectum for the patient 16 was -7.17 Gy less than SS. For this case, there was huge amount of gas inside the rectum in the second fraction compared to the first fraction (the volume difference was 127 cc). Table 4 and Figure 4 show that the number of outlier data of the bladder for each criterion (mild outlier, moderate outliers, extreme outlier) is greater than that for the rectum. One reason might be that the interfraction volume differences of the bladder were larger than those for the rectum.

In some patients, the accumulated dose of $D_{0.1cc}$ for the bladder was observed to be greater than by the SS approach (supplements Figure 1). It is possibly due to the uncertainty of the hybrid-based DIR algorithm. $D_{0.1cc}$ of the bladder in these cases was close to the source where the dose gradient is high. On the other hand, the hybrid-based DIR algorithm expands the small bladder contour (second or third Fr) to the large bladder contour (first Fr), causing increased $D_{0.1cc}$ due to the steep gradient dose region [31].

One disadvantage of DICE and Jaccard is that it does not guarantee voxel-to-voxel overlapping inside the contour, though it does not affect the results [18]. Wachter-Gerstner *et al.* [36] reported that D_{2cc} calculated from DVH of the external contour is sufficient to estimate bladder and rectum wall doses as both are hollow organs.

There were some limitations in our study including uncertainty in delineation of organs in CT images, inter-observer contouring uncertainty on OAR, uncertainty of hybrid-based DIR, and assuming that the entire prescription EBRT dose was delivered to OARs. MRI-based BT decreases uncertainty of delineation organs due to better soft tissue contrast. Hellebust *et al.* [37] found 5-8% inter-observer variability using MRI images compared to 10-11% obtained by Saarnak *et al.* [38] using CT images for

bladder and rectum. It is hard to distinguish to what extent dose differences between two approaches are related to algorithm uncertainty. Therefore all of the results calculated with the DIR algorithm contain DIR uncertainty.

In this study, we investigated DIR only among BT fractions but it is more challenging to combine EBRT and HDR-BT to assess DA due to the presence of the applicator in BT. Also, we did not consider interfraction motion in EBRT, intrafraction motion in BT, or the effect of DIR in radiobiological modeling. Further research is required to confirm our results and especially evaluate the DIR algorithm with ground truth data the same as for the physical phantom. We plan to investigate the feasibility of predicting rectal toxicity with deformable accumulated and dose map features for locally advanced cervical cancer.

Conclusions

We evaluated the hybrid-based DIR algorithm using a virtual phantom and demonstrated low registration uncertainty for both contour matching and dose mapping. Our results demonstrated that differences in DA between DIR-based and SS approaches were statistically significant for both bladder and rectum. Also, the results indicated that the hybrid-based DIR algorithm has potential for DA between brachytherapy fractions.

Acknowledgments

This work was supported by Iran University of Medical Sciences, grant number 32340.

Disclosure

The authors report no conflict of interest.

References

- Dimopoulos JC, Petrow P, Tanderup K et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (IV): Basic principles and parameters for MR imaging within the frame of image based adaptive cervix cancer brachytherapy. *Radiother Oncol* 2012; 103: 113-122.
- Haie-Meder C, Pötter R, Van Limbergen E et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (I): concepts and terms in 3D image based 3D treatment planning in cervix cancer brachytherapy with emphasis on MRI assessment of GTV and CTV. *Radiother Oncol* 2005; 74: 235-245.
- Hellebust TP, Kirisits C, Berger D et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group: considerations and pitfalls in commissioning and applicator reconstruction in 3D image-based treatment planning of cervix cancer brachytherapy. *Radiother Oncol* 2010; 96: 153-160.
- Pötter R, Haie-Meder C, Van Limbergen E et al. Recommendations from gynaecological (GYN) GEC ESTRO working group (II): Concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy – 3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. *Radiother Oncol* 2006; 78: 67-77.
- Viswanathan AN, Beriwal S, De Los Santos JF et al. American Brachytherapy Society consensus guidelines for locally advanced carcinoma of the cervix. Part II: high-dose-rate brachytherapy. *Brachytherapy* 2012; 11: 47-52.
- Viswanathan AN, Thomadsen B, and American Brachytherapy Society Cervical Cancer Recommendations Committee; American Brachytherapy Society. American Brachytherapy Society consensus guidelines for locally advanced carcinoma of the cervix. Part I: general principles. *Brachytherapy* 2012; 11: 33-46.
- Hellebust TP, Dale E, Skjønberg A et al. Inter fraction variations in rectum and bladder volumes and dose distributions during high dose rate brachytherapy treatment of the uterine cervix investigated by repetitive CT-examinations. *Radiother Oncol* 2001; 60: 273-280.
- Kobayashi K, Murakami N, Wakita A et al. Dosimetric variations due to interfraction organ deformation in cervical cancer brachytherapy. *Radiother Oncol* 2015; 117: 555-558.
- Li S, Liu Z, Chan P et al. Feasibility of composing external beam radiotherapy and HDR brachytherapy of cervical cancer using deformable image registration. *Int J Radiat Oncol Biol Phys* 2011; 81: S464-S465.
- Millar LB, Teo BK, Lin L. Using deformable image registration to assess the accuracy of cumulative external beam and intracavitary brachytherapy ICRU point doses in cervical cancer. *Int J Radiat Oncol Biol Phys* 2010; 78: S411-S412.
- Chapman CH, Polan D, Jolly S et al. Deformable image registration improves contouring accuracy in magnetic resonance imaging-based cervical brachytherapy. *Int J Radiat Oncol Biol Phys* 2016; 96: E669.
- Siavashpour Z, Aghamiri MR, Jaberi R et al. Evaluating the utility of “3D Slicer” as a fast and independent tool to assess intrafractional organ dose variations in gynecological brachytherapy. *Brachytherapy* 2016; 15: 514-523.
- Christensen GE, Carlson B, Chao KS et al. Image-based dose planning of intracavitary brachytherapy: registration of serial-imaging studies using deformable anatomic templates. *Int J Radiat Oncol Biol Phys* 2001; 51: 227-243.
- Flower E, Do V, Sykes J et al. Deformable image registration for cervical cancer brachytherapy dose accumulation: Organ at risk dose-volume histogram parameter reproducibility and anatomic position stability. *Brachytherapy* 2017; 16: 387-392.
- Chapman CH, Polan D, Vineberg K et al. Deformable image registration-based contour propagation yields clinically acceptable plans for MRI-based cervical cancer brachytherapy planning. *Brachytherapy* 2018; 17: 360-367.
- van Heerden LE, van Wieringen N, Koedooder K et al. Dose warping uncertainties for the accumulated rectal wall dose in cervical cancer brachytherapy. *Brachytherapy* 2018; 17: 449-455.
- Kadoya N, Miyasaka Y, Yamamoto T et al. Evaluation of rectum and bladder dose accumulation from external beam radiotherapy and brachytherapy for cervical cancer using two different deformable image registration techniques. *J Radiat Res* 2017; 58: 720-728.
- van Heerden LE, Houweling AC, Koedooder K et al. Structure-based deformable image registration: Added value for dose accumulation of external beam radiotherapy and brachytherapy in cervical cancer. *Radiother Oncol* 2017; 123: 319-324.
- Brock KK. Image processing in radiation therapy. CRC Press, 2013.
- Hayashi K, Isohashi F, Akino Y et al. Estimation of the total rectal dose of radical external beam and intracavitary radiotherapy for uterine cervical cancer using the deformable image registration method. *J Radiat Res* 2015; 56: 546-552.
- Reniers B, Janssens G, Orban de Xivry J et al. Dose distribution for gynecological brachytherapy with dose accumulation between insertions: Feasibility study. *Brachytherapy* 2016; 15: 504-513.

22. Varadhan R, Karangelis G, Krishnan K et al. A framework for deformable image registration validation in radiotherapy clinical applications. *J Appl Clin Med Phys* 2013; 14: 192-213.
23. Bookstein F.L. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1989; 11: 567-585.
24. Wendland H. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics* 1995; 4: 389-396.
25. Albuquerque K, Hrycushko BA, Harkenrider MM et al. Compendium of fractionation choices for gynecologic HDR brachytherapy – An American Brachytherapy Society Task Group Report. *Brachytherapy* 2019; 18: 429-436.
26. Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945; 26: 297-302.
27. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1993; 15: 850-863.
28. Olteanu LA, Madani I, De Neve W et al. Evaluation of deformable image coregistration in adaptive dose painting by numbers for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2012; 83: 696-703.
29. Xu Z, Lee CP, Heinrich MP et al. Evaluation of six registration methods for the human abdomen on clinically acquired CT. *IEEE Trans Biomed Eng* 2016; 63: 1563-1572.
30. Sabater S, Andres I, Sevellano M et al. Dose accumulation during vaginal cuff brachytherapy based on rigid/deformable registration vs. single plan addition. *Brachytherapy* 2014; 13: 343-351.
31. Jamema SV, Mahantshetty U, Andersen E et al. Uncertainties of deformable image registration for dose accumulation of high-dose regions in bladder and rectum in locally advanced cervical cancer. *Brachytherapy* 2015; 14: 953-962.
32. Andersen ES, Noe KØ, Sørensen TS et al. Simple DVH parameter addition as compared to deformable registration for bladder dose accumulation in cervix cancer brachytherapy. *Radiother Oncol* 2013; 107: 52-57.
33. Nesvacil N, Tanderup K, Lindegaard JC et al. Can reduction of uncertainties in cervix cancer brachytherapy potentially improve clinical outcome? *Radiother Oncol* 2016; 120: 390-396.
34. Kirisits C, Rivard MJ, Baltas D et al. Review of clinical brachytherapy uncertainties: analysis guidelines of GEC-ESTRO and the AAPM. *Radiother Oncol* 2014; 110: 199-212.
35. Siavashpour Z, Aghamiri MR, Jaber R et al. Optimum organ volume ranges for organs at risk dose in cervical cancer intracavitary brachytherapy. *J Contemp Brachytherapy* 2016; 8: 135-142.
36. Wachter-Gerstner N, Wachter S, Reinstadler E et al. Bladder and rectum dose defined from MRI based treatment planning for cervix cancer brachytherapy: comparison of dose-volume histograms for organ contours and organ wall, comparison with ICRU rectum and bladder reference point. *Radiother Oncol* 2003; 68: 269-276.
37. Hellebust TP, Tanderup K, Lervåg C et al. Dosimetric impact of interobserver variability in MRI-based delineation for cervical cancer brachytherapy. *Radiother Oncol* 2013; 107: 13-19.
38. Saarnak AE, Boersma M, van Bunnigen BN et al. Inter-observer variation in delineation of bladder and rectum contours for brachytherapy of cervical cancer. *Radiother Oncol* 2000; 56: 37-42.