

NOTE

## Deformable anthropomorphic pelvis phantom for dose accumulation verification

To cite this article: Yun Ming Wong *et al* 2024 *Phys. Med. Biol.* **69** 12NT01

View the [article online](#) for updates and enhancements.

### You may also like

- [Crystallinity characterization of white matter in the human brain](#)  
Erin G Teich, Matthew Cieslak, Barry Giesbrecht et al.
- [Automatic re-contouring in 4D radiotherapy](#)  
Weiguo Lu, Gustavo H Olivera, Quan Chen et al.
- [Efficient voxel navigation for proton therapy dose calculation in TOPAS and Geant4](#)  
J Schümann, H Paganetti, J Shin et al.



## NOTE

## Deformable anthropomorphic pelvis phantom for dose accumulation verification

RECEIVED  
25 February 2024REVISED  
30 April 2024ACCEPTED FOR PUBLICATION  
31 May 2024PUBLISHED  
11 June 2024Yun Ming Wong<sup>1,8</sup> , Calvin Wei Yang Koh<sup>2,7</sup>, Kah Seng Lew<sup>1,2</sup>, Clifford Ghee Ann Chua<sup>2,7</sup>, Ping Lin Yeap<sup>2,4</sup>, Ee Teng Zhang<sup>5,6</sup>, Ashley Li Kuan Ong<sup>2,7</sup>, Jeffrey Kit Loong Tuan<sup>2,7</sup>, Bing Feng Ng<sup>5</sup>, Wen Siang Lew<sup>1</sup>, James Cheow Lei Lee<sup>1,2</sup> and Hong Qi Tan<sup>1,2,3,7,\*</sup> <sup>1</sup> Division of Physics and Applied Physics, Nanyang Technological University, Singapore, Singapore<sup>2</sup> Division of Radiation Oncology, National Cancer Centre Singapore, Singapore, Singapore<sup>3</sup> Oncology Academic Clinical Programme, Duke-NUS Medical School, Singapore, Singapore<sup>4</sup> Department of Oncology, University of Cambridge, Cambridge, United Kingdom<sup>5</sup> School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore, Singapore<sup>6</sup> Singapore Centre for 3D Printing, Nanyang Technological University, Singapore, Singapore<sup>7</sup> Ai3 Lab, National Cancer Centre Singapore, Singapore, Singapore<sup>8</sup> First authors.

\* Author to whom any correspondence should be addressed.

E-mail: [tan.hong.qi@nccs.com.sg](mailto:tan.hong.qi@nccs.com.sg)**Keywords:** deformable image registration, adaptive radiotherapy, dice similarity coefficient, target registration error, dose accumulation, 3D printing, deformable anthropomorphic phantomSupplementary material for this article is available [online](#)**Abstract**

**Objective.** The validation of deformable image registration (DIR) for contour propagation is often done using contour-based metrics. Meanwhile, dose accumulation requires evaluation of voxel mapping accuracy, which might not be accurately represented by contour-based metrics. By fabricating a deformable anthropomorphic pelvis phantom, we aim to (1) quantify the voxel mapping accuracy for various deformation scenarios, in high- and low-contrast regions, and (2) identify any correlation between dice similarity coefficient (DSC), a commonly used contour-based metric, and the voxel mapping accuracy for each organ. **Approach.** Four organs, i.e. pelvic bone, prostate, bladder and rectum (PBR), were 3D printed using PLA and a Polyjet digital material, and assembled. The latter three were implanted with glass bead and CT markers within or on their surfaces. Four deformation scenarios were simulated by varying the bladder and rectum volumes. For each scenario, nine DIRs with different parameters were performed on RayStation v10B. The voxel mapping accuracy was quantified by finding the discrepancy between true and mapped marker positions, termed the target registration error (TRE). Pearson correlation test was done between the DSC and mean TRE for each organ. **Main results.** For the first time, we fabricated a deformable phantom purely from 3D printing, which successfully reproduced realistic anatomical deformations. Overall, the voxel mapping accuracy dropped with increasing deformation magnitude, but improved when more organs were used to guide the DIR or limit the registration region. DSC was found to be a good indicator of voxel mapping accuracy for prostate and rectum, but a comparatively poorer one for bladder.  $DSC > 0.85/0.90$  was established as the threshold of mean TRE  $\leq 0.3$  cm for rectum/prostate. For bladder, extra metrics in addition to DSC should be considered. **Significance.** This work presented a 3D printed phantom, which enabled quantification of voxel mapping accuracy and evaluation of correlation between DSC and voxel mapping accuracy.

**1. Introduction**

Adaptive radiotherapy (ART), a paradigm which allows adaptation of treatment plans based on patients' latest anatomy (Yan *et al* 1997), is recognized for ensuring an optimal target coverage and minimal dose to organs at risk (OARs). This is especially beneficial for cancer sites which involves considerable anatomical

changes throughout the radiotherapy course, such as head and neck (H&N) cancer (Schwartz *et al* 2013, Buciuman and Marcu 2022, Avkshtol *et al* 2023), lung cancer (Agarwal *et al* 2020, Møller *et al* 2022, Hoppen *et al* 2023), prostate cancer (Nijkamp *et al* 2008, Christiansen *et al* 2022, Waters *et al* 2024), etc.

In this context, deformable image registration (DIR) plays an important role by revealing information about the transformation between images taken at different time point during the treatment course. There are three major DIR applications in ART, namely, contour propagation, dose accumulation (Rigaud *et al* 2019), and synthetic CT generation for dose calculation (Irmak *et al* 2020, Sheikh *et al* 2022). All these require a thorough DIR validation to verify the registration accuracy before they can be translated for clinical use. The first case is undoubtedly more straightforward as it can easily be evaluated by contour-based metrics, which is one of the most commonly used metrics. Some examples include dice similarity coefficient (DSC), mean distance to agreement, and Hausdorff distance. Meanwhile, the credibility of synthetic CT is often evaluated through gamma analysis of dose distributions (Thing *et al* 2022, Chang *et al* 2023).

In comparison, DIR validation for dose accumulation calls for more caution as the accuracy here matters not only for the organ boundaries but also for the volumes within. While contour-based metrics seem to offer an efficient way for DIR validation, they often do not confirm an accurate registration of the volume within the contour (Varadhan *et al* 2013). A former study by Rohlfing (2012) demonstrated the unreliability of tissue overlap measures to identify inaccurate registrations. Only sufficiently small landmarks were able to distinguish between reasonable and poor registrations. Using a two-dimensional deformable phantom with nonradiopaque markers on the phantom surface, Kirby *et al* (2013) have also shown that a DIR algorithm with the highest DSC could surprisingly give the largest spatial error. These findings suggested that landmarks might be a more reliable means for DIR validation, in terms of voxel mapping accuracy which is of particular significance for dose accumulation application.

That being said, landmark techniques are highly dependent on image quality and contrast (Paganelli *et al* 2018). Landmark identification require distinct features that are clearly visible on imaging scans, for example, the vallecule and philtrum for H&N cancer (Rigaud *et al* 2015), and vessel bifurcations for lung cancer (Castillo *et al* 2009) as well as liver cancer (Zhang 2023). Due to the lack of distinct features in the prostate, evaluation of registration accuracy using landmarks is uncommon.

In instances where precise motion tracking or prostate matching is required, fiducial markers may be implanted into the prostate (Kupelian *et al* 2005, Cao *et al* 2022). Similar ideas have been adopted in a few deformable pelvis phantoms, where markers were implanted within the prostate (Wang *et al* 2005) or on the surface of several organs (Bohoudi *et al* 2019) for DIR validation. While these studies have attempted to quantify the 'landmark-based' registration errors, none has performed an organ-wise correlation analysis between the contour-based metrics and the landmark techniques.

To this end, we fabricated a deformable anthropomorphic pelvis phantom using 3D printing materials, with markers placed on or within the organs. Other than allowing an end-to-end test (Brock *et al* 2017), a deformable anthropomorphic phantom poses the advantage of realistic anatomical deformations compared to simpler geometric phantom. Niebuhr *et al* have previously constructed such a pelvis phantom—the ADAM-pelvis phantom—by 3D printing the organ shapes based on patient images, and formed the organs using surrogate materials such as gypsum and silicone (2019). Here, we proposed a rapid and low-cost alternative for customizing a deformable phantom through 3D printing, by having a contoured CT scan as the main 'ingredient'. The great variety of 3D printing materials available, ranging from rigid to elastic options, is also a pivotal element for the creation of organ models with realistic deformations.

In our analysis, three primary organs were included: the target (prostate) and two OARs (bladder, rectum). Markers were attached throughout the entire organs walls to quantify the target registration errors (TRE). They were designed to be easily attachable or detachable from the organ walls, making it possible to adjust the marker positions as necessary. Given the observed enhancement in DIR performance with the addition of markers (Varadhan *et al* 2016), our study incorporated a marker overriding step before the DIR process. By simulating various deformation scenarios and modifying the DIR parameters, we hope to achieve two main objectives: (1) to assess the DIR voxel mapping accuracy using markers for a range of deformation magnitudes, and (2) to identify any correlation between contour-based metric and voxel mapping accuracy. Additionally, we will examine the mechanical properties of the 3D printing materials used, to shed light on their suitability for our deformable phantom (supplementary material).

## 2. Methods

### 2.1. Phantom fabrication

#### 2.1.1. 3D printing

To reproduce an anthropomorphic phantom for prostate cancer patient, the main organs of concern are the prostate, bladder, rectum and pelvic bone. These organ contours were extracted from the CT scan of one

patient (approved by the SingHealth Centralized Institutional Review Board) into .STL files. Further processing was then done on the PBR files as follows. All three organs were made hollow with a wall thickness of 2 mm. 15 mm and 8 mm holes were created at the top and bottom of the prostate, respectively. Two tubes (inner diameter (id) 9 mm, outer diameter (od) 13 mm) were added to the top of the bladder to simulate the ureters and one tube (id 4 mm, od 8 mm) was added to the bottom as the urethra. Lastly, the rectum was modified such that it acts as an open channel which allows smooth passage of fluids. Other than representing a real-life situation, these openings are also important for easy removal of support materials after printing. The prostate, bladder and rectum were 3D printed using a shore 50A PolyJet digital material which is a mix of Agilus30 Clear (FLX935) and Vero PureWhite (RGD 837), whereas the pelvic bone was printed using PolyMax™ PLA with an infill density of 10%. MakerBot CloudPrint (Ultimaker B.V., Geldermalsen, The Netherlands) was used for slicing. Due to the size limitation of a MakerBot Method X (Ultimaker B.V., Geldermalsen, The Netherlands) 3D printer in printing a human-sized pelvic bone, the bone was split into four quadrants and fixed together after printing.

### 2.1.2. Marker fixation

To fill up the hollow prostate model, an inexpensive and readily available ingredient was used—agar. 2 g of agar powder was mixed with 30 ml of water in a beaker and heated at low heat until it started to boil. This concentration was used to create a medium with CT number between 30 and 50 HU (Niebuhr *et al* 2019), so that its rigidity resembles a real prostate. The agar was added into the prostate model and 2 mm glass bead markers were inserted at various positions around the medium. After the agar cooled, bead and CT markers were attached uniformly across the wall of the prostate using silicone sealant. The markers were attached in a similar manner for bladder and rectum (figure 1).

### 2.1.3. Phantom assembly

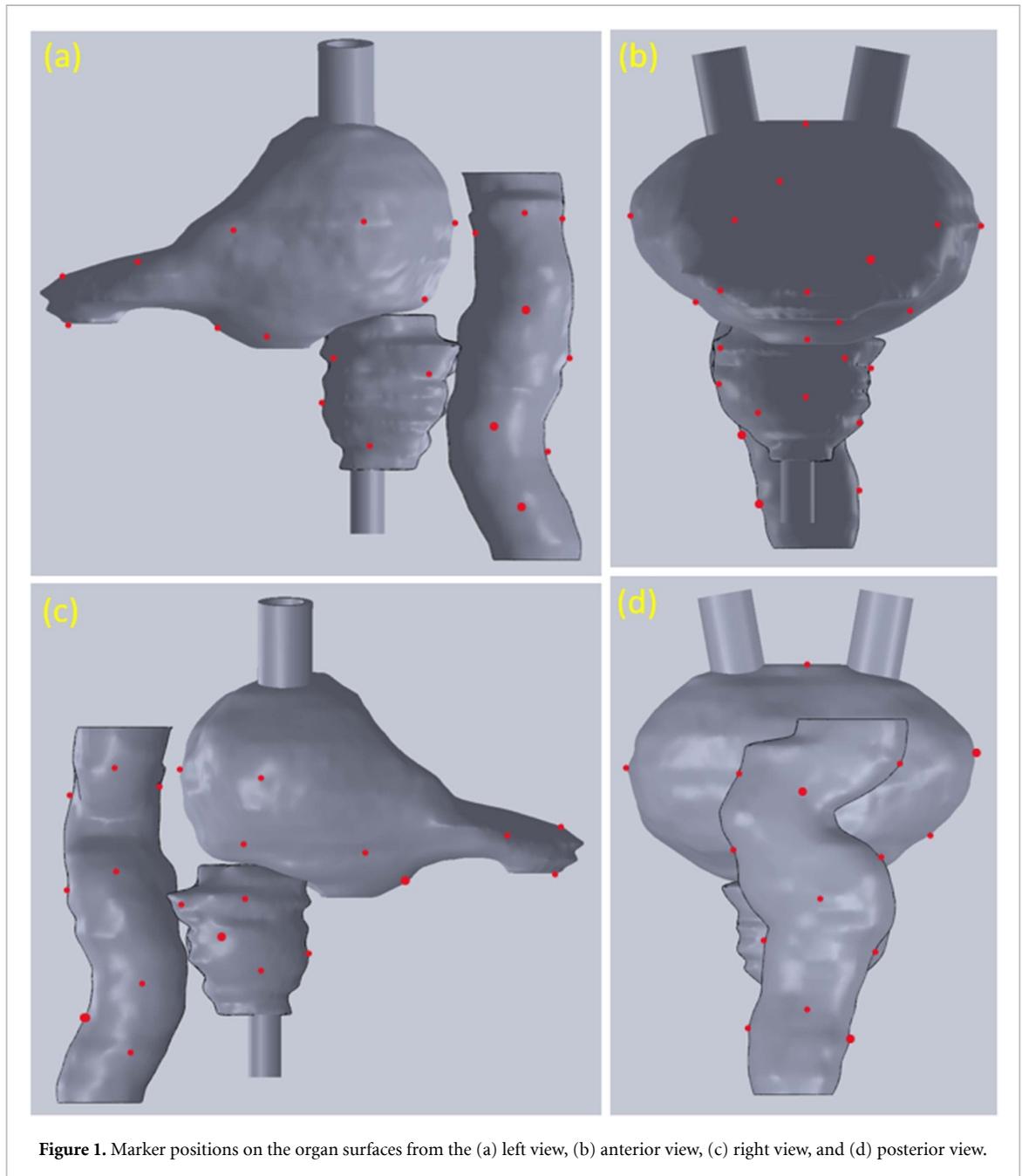
The pelvic bone and rectum were first fixed in place in a rectangular acrylic case by pouring a layer of silicone at the bottom. The bladder and prostate were combined through the urethra (which has been blocked), and the former was attached to the pubic tubercle of the pelvic bone while the latter was attached to the anterior wall of the rectum. The filling mechanisms of bladder and rectum were done in a similar way to the aforementioned ADAM-pelvis phantom study (Niebuhr *et al* 2019): Plastic tubes were connected to the ureters at one end and syringes at the other end to vary the amount of water for simulation of bladder volume change; Rectal balloon (Boston Scientific, MA, USA) was placed inside the rectum and was also connected to a syringe to expand the rectum with air.

## 2.2. Proof of concept study

To mimic a realistic human body where tissues surround the organs, the assembled phantom was filled with water (which is considered a closer tissue-equivalent material), to a level that is sufficient to submerge the pelvic bone completely (figure 2(a)). A CT scan was taken using Siemens SOMATOM X.cite (Siemens Healthineers, Forchheim, Germany) without any deformation (CT 1). Four deformation scenarios were then simulated on the bladder and prostate (table 1). The deformation volumes were decided by taking three patients' organ volume changes as an empirical reference. For each deformation scenario, a CT scan was taken (CT 2–5). This resulted in five CT scans in total, which were imported into RayStation v10B (RaySearch Laboratories AB, Stockholm, Sweden). Each marker was designated as a point of interest (POI) on all scans, and the marker movements were determined by finding the Euclidean distance between the corresponding POIs on CT 1 and CT 2–5.

The same five scans were also imported into 3D Slicer ('3D Slicer image computing platform' *n.d.*) to override the CT number according to the mean CT number obtained from the actual patient CT scan (table 2). This step was done for two main purposes: (1) to increase the contrast on the images as most regions on the original scans consisted of water, and (2) to remove external objects, such as tubes and markers, from the scans. Examples of one slice of CT 1 before and after overriding are shown in figures 2(b) and (c), respectively. These overridden images were subsequently imported into RayStation. Using CT 1 as the reference image and the CT 2–5 as the target images, nine DIRs were performed for each pair of images. All DIRs were hybrid intensity and structure based deformable registrations optimized using correlation coefficient as the similarity measure and 0.25 cm as the final resolution. They differed in terms of the controlling ROI (CROI) and focus ROI (FROI) chosen. The first DIR was not guided by any CROI or FROI, the second and third used bladder, the fourth and fifth used prostate, the sixth and seventh used rectum, and the eighth and ninth used PBR as CROI or FROI. This gave rise to nine DIRs with different registration qualities.

From the initial positions of markers on CT 1, the marker positions after deformation were determined by mapping the points based on both rigid registration and DIR, using a built-in scripting function in



**Figure 1.** Marker positions on the organ surfaces from the (a) left view, (b) anterior view, (c) right view, and (d) posterior view.

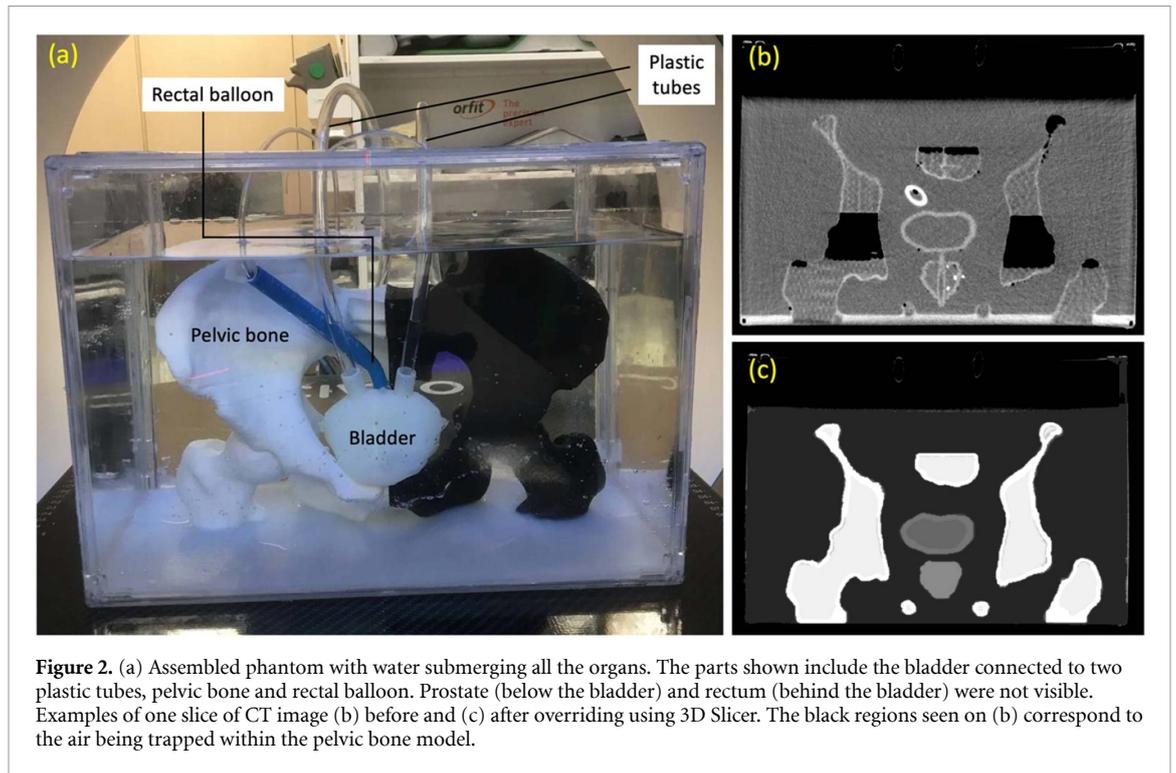
RayStation. The distance between the mapped points and original points on CT 2–5 (ground truth) would be TRE. To better analyze the effect of CROI or FROI on the registration accuracy of each organ, the error percentages with respect to the mean marker movements were calculated as shown in equation (1). The DSC and mean TRE of PBR were obtained for each DIR, after which a Pearson correlation analysis was done to test the null hypothesis that there is no correlation between the two metrics. A two tailed  $p$ -value of 0.05 marked the significance of the test.

$$\text{Error percentage} = \frac{\text{Mean TRE}}{\text{Mean marker movement}} \times 100\%. \quad (1)$$

### 3. Results

The mean marker movement and TRE for each deformation scenario and each DIR are tabulated in table 3. A larger marker movement generally entailed a larger TRE. Besides, an improvement in registration accuracy was observed when more organs were used as CROI or FROI.

The marker movement and TRE for each marker in the first deformation scenario are plotted in figure 3. Figure 3(a) shows the results from DIR without guidance from any CROI/FROI whereas figure 3(b) shows

**Table 1.** Four deformation scenarios simulated.

CT	Rectum expansion (ml)	Bladder expansion (ml)
1	—	—
2	25	30
3	25	60
4	50	30
5	50	60

**Table 2.** Details of CT number overriding based on the patient CT scan, using 3D Slicer.

Organ/Region	Overridden CT number
Pelvic bone	<ul style="list-style-type: none"> <li>• Outer (2 mm): 641.77 HU</li> <li>• Inner: 190.11 HU</li> </ul>
Prostate	50 HU
Bladder	<ul style="list-style-type: none"> <li>• Outer (3 mm): 21.12 HU</li> <li>• Inner: 0 HU</li> </ul>
Rectum	<ul style="list-style-type: none"> <li>• Outer (2 mm): 24.05 HU</li> <li>• Inner: 27.76 HU</li> </ul>
Remaining region (adipose)	−89.46 HU

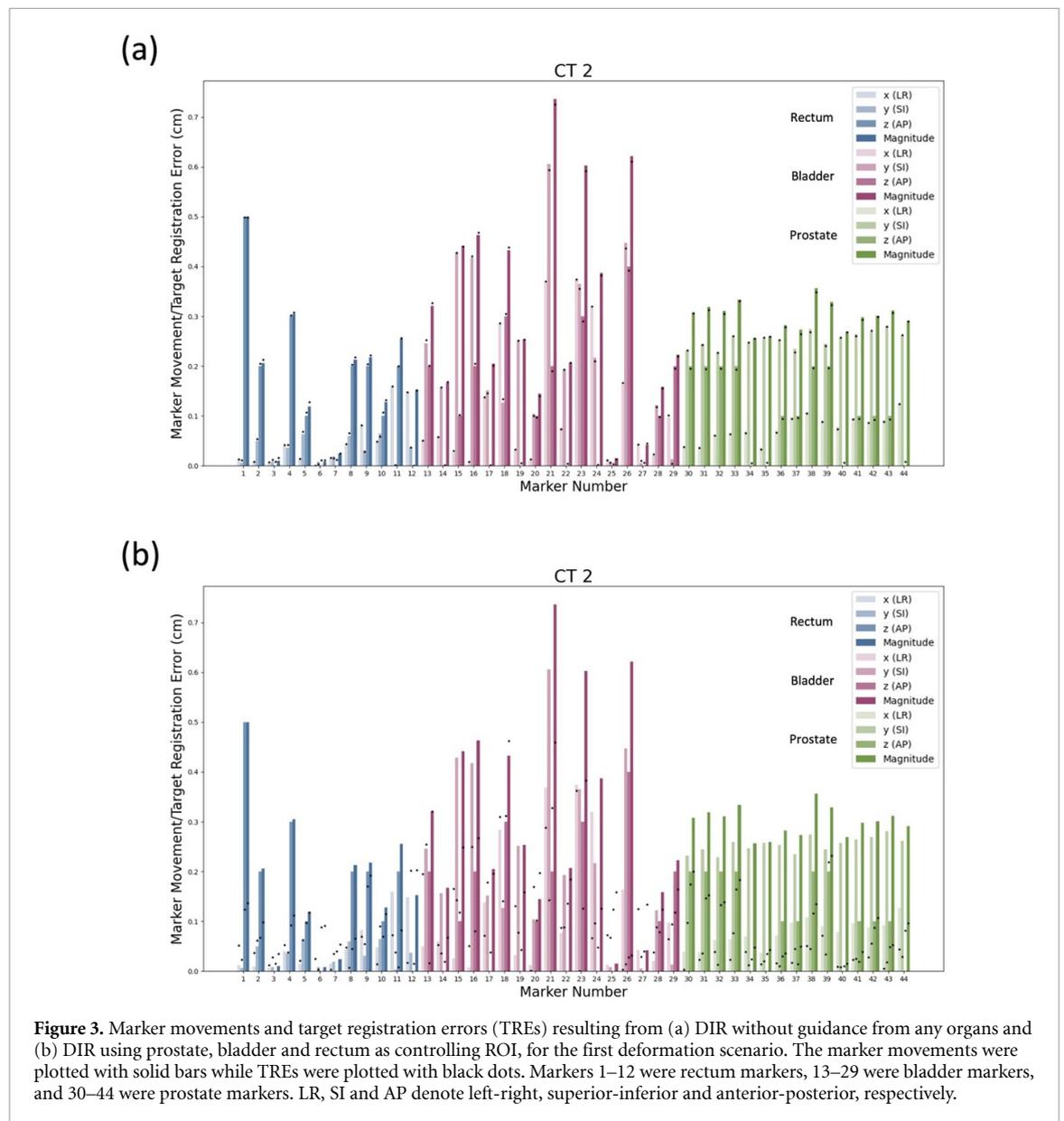
the results from DIR using PBR as CROI. The marker movements (depicted by the solid color bars) were the same in both cases since they originated from the same deformation scenario, but the TREs (depicted by the black dots) varied. It is obvious that the former case had TREs which were similar in magnitude with the marker movements. In the latter case, most markers (other than those with smaller movements) achieved a reduction in TREs. This is especially evident for markers with larger movements such as marker 21 and 26.

Tables 4 and S1 in supplementary material show the error percentages for each organ, DIR and deformation scenario. Looking across the rows, it is apparent that the error percentages were relatively smaller for the organ(s) used as either CROI or FROI in each DIR (highlighted in bold in both the tables).

**Table 3.** Mean marker movements and target registration errors ( $\pm$  standard errors) for each DIR and deformation scenario.

	CT	CT 2	CT 3	CT 4	CT 5
Mean marker movement (cm)		0.274 $\pm$ 0.024	0.362 $\pm$ 0.028	0.622 $\pm$ 0.053	0.641 $\pm$ 0.047
	NROI	0.273 $\pm$ 0.023	0.357 $\pm$ 0.028	0.619 $\pm$ 0.053	0.640 $\pm$ 0.048
	bCROI	0.219 $\pm$ 0.017	0.282 $\pm$ 0.022	0.481 $\pm$ 0.042	0.469 $\pm$ 0.044
	bFROI	0.193 $\pm$ 0.016	0.243 $\pm$ 0.019	0.395 $\pm$ 0.035	0.364 $\pm$ 0.032
	pCROI	0.214 $\pm$ 0.025	0.308 $\pm$ 0.032	0.444 $\pm$ 0.053	0.480 $\pm$ 0.053
Mean target registration error (cm)	pFROI	0.214 $\pm$ 0.025	0.301 $\pm$ 0.032	0.412 $\pm$ 0.047	0.431 $\pm$ 0.044
	rCROI	0.249 $\pm$ 0.024	0.326 $\pm$ 0.030	0.527 $\pm$ 0.056	0.563 $\pm$ 0.050
	rFROI	0.218 $\pm$ 0.025	0.284 $\pm$ 0.030	0.380 $\pm$ 0.043	0.360 $\pm$ 0.046
	PBR_CROI	0.148 $\pm$ 0.016	0.203 $\pm$ 0.021	0.261 $\pm$ 0.032	0.309 $\pm$ 0.038
	PBR_FROI	0.165 $\pm$ 0.015	0.204 $\pm$ 0.017	0.294 $\pm$ 0.027	0.306 $\pm$ 0.031

Abbreviations: NROI—DIR without CROI/FROI; bCROI/bFROI—DIR using bladder as CROI/FROI; pCROI/pFROI—DIR using prostate as CROI/FROI; rCROI/rFROI—DIR using rectum as CROI/FROI; PBR\_CROI/PBR\_FROI—DIR using all three organs as CROI/FROI.

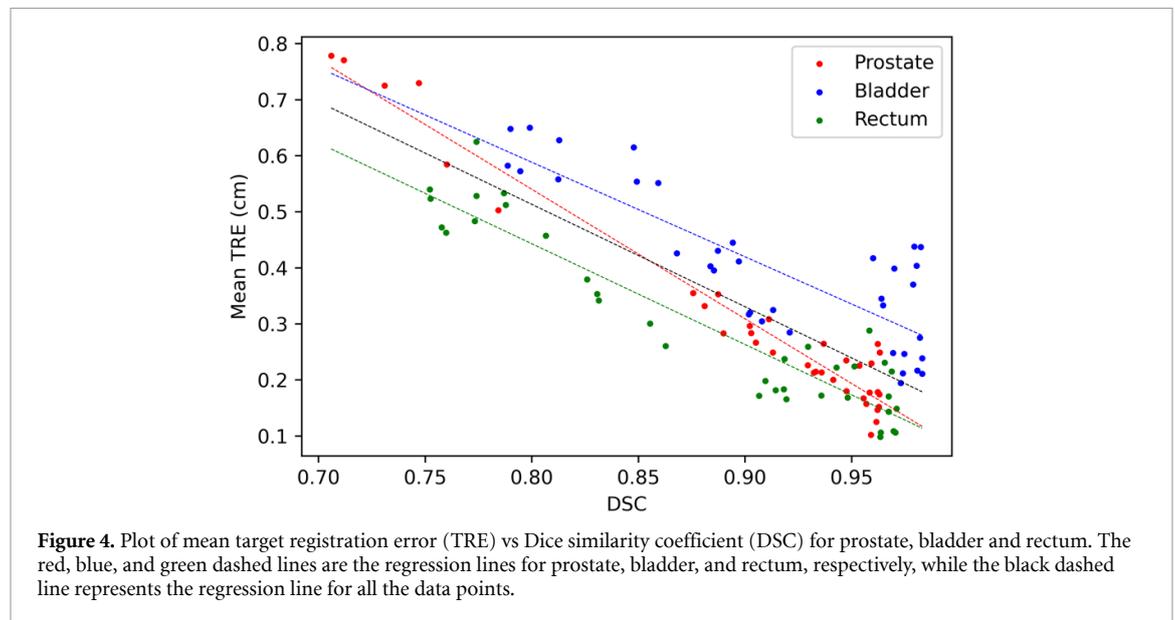


**Figure 3.** Marker movements and target registration errors (TREs) resulting from (a) DIR without guidance from any organs and (b) DIR using prostate, bladder and rectum as controlling ROI, for the first deformation scenario. The marker movements were plotted with solid bars while TREs were plotted with black dots. Markers 1–12 were rectum markers, 13–29 were bladder markers, and 30–44 were prostate markers. LR, SI and AP denote left-right, superior-inferior and anterior-posterior, respectively.

**Table 4.** Mean marker movements and error percentages relative to the marker movement for each organ and DIR, in the first deformation scenario. The values corresponding to the organs used as CROI/FROI were highlighted in bold.

CT	CT 2		
	P	B	R
Mean marker movement (cm)	0.300	0.319	0.179
NROI	98.7	99.3	101.7
bCROI	88.9	<b>66.1</b>	96.1
bFROI	71.0	<b>60.9</b>	92.6
pCROI	<b>41.7</b>	95.5	110.8
pFROI	<b>52.4</b>	89.3	102.6
rCROI	94.4	100.4	<b>59.3</b>
rFROI	66.7	101.8	<b>59.4</b>
PBR_CROI	<b>34.0</b>	<b>67.9</b>	<b>60.9</b>
PBR_FROI	<b>55.7</b>	<b>66.3</b>	<b>55.1</b>

Abbreviations: P—Prostate; B—Bladder; R—Rectum; NROI—DIR without CROI/FROI; bCROI/bFROI—DIR using bladder as CROI/FROI; pCROI/pFROI—DIR using prostate as CROI/FROI; rCROI/rFROI—DIR using rectum as CROI/FROI; PBR\_CROI/PBR\_FROI—DIR using all three organs as CROI/FROI.



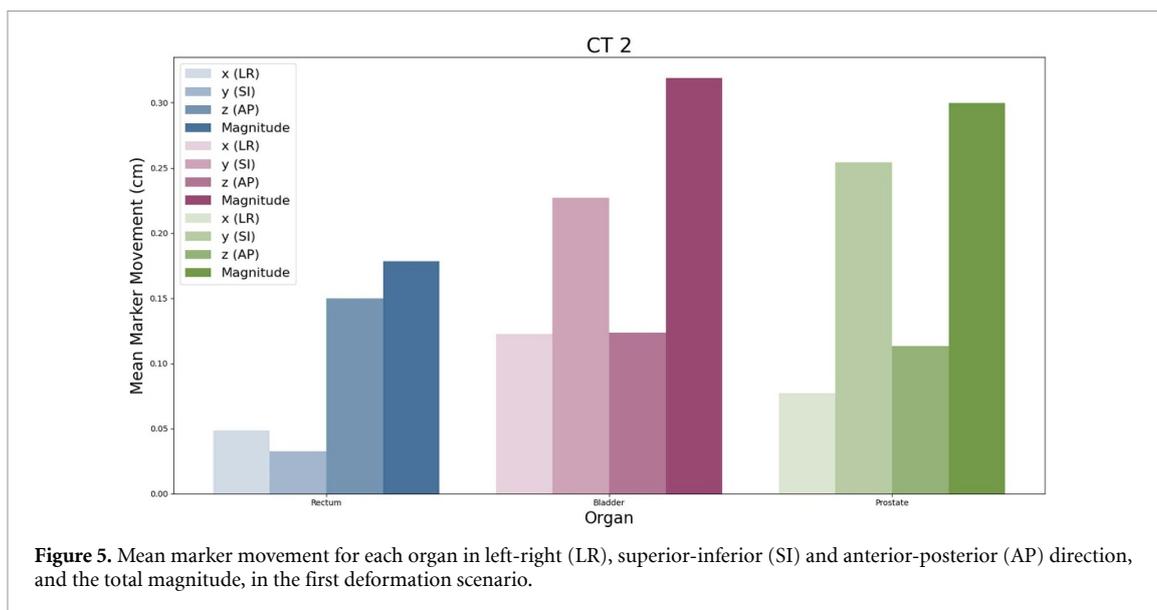
**Figure 4.** Plot of mean target registration error (TRE) vs Dice similarity coefficient (DSC) for prostate, bladder and rectum. The red, blue, and green dashed lines are the regression lines for prostate, bladder, and rectum, respectively, while the black dashed line represents the regression line for all the data points.

**Table 5.** Pearson correlation coefficient between DSC and mean TRE and its associated p-value for prostate, bladder, rectum, and all three organs together.

	Prostate	Bladder	Rectum	All
Pearson $r$	-0.9772	-0.8286	-0.9358	-0.8347
Pearson $p$	<0.01	<0.01	<0.01	<0.01

Interestingly, the homogenous medium within the prostate did not give rise to a terribly erroneous voxel mapping. In the extreme case (CT 4) where the mean marker movement for the markers inside prostate was close to 0.75 cm, the mean TRE could be as small as 0.14 cm (with DIR applying PBR as CROI).

Figure 4 illustrates the relationship between the DSC and mean TRE of each organ, resulting from different deformation scenarios and DIRs. Pearson correlation tests revealed a statistically significant ( $P \ll 0.05$ ) very strong negative correlation ( $r \leq -0.80$ ) between the DSC and mean TRE for prostate, bladder, rectum, and all three organs considered together (table 5).



**Figure 5.** Mean marker movement for each organ in left-right (LR), superior-inferior (SI) and anterior-posterior (AP) direction, and the total magnitude, in the first deformation scenario.

#### 4. Discussion

In this work, we fabricated a deformable anthropomorphic phantom using 3D printing materials, and demonstrated its feasibility for DIR quality assurance (QA) in terms of voxel mapping accuracy through quantification of TRE.

Past studies (Crook *et al* 1995, Roeske *et al* 1995, Dawson *et al* 1998, Nederveen *et al* 2002, Britton *et al* 2005, Litzenberg *et al* 2006, Kotte *et al* 2007, Pang *et al* 2018) have shown that the prostate motion is mainly in the anterior-posterior (AP) and superior-inferior (SI) directions (corresponding to  $z$ - and  $y$ -direction, respectively, in our study). By inducing deformation of bladder and rectum in our phantom, we managed to reproduce a realistic prostate deformation whereby SI motion dominates (figures 5 and S3–5 in supplementary material).

One may argue that the Polyjet digital material has a tensile modulus that is much lower than a human bladder and rectum (0.25 MPa (Dahms *et al* 1998) and 5.18 MPa (Christensen *et al* 2015), respectively). As a matter of fact, it is possible to tune the tensile modulus by modifying the proportion of Agilus30 Clear to Vero PureWhite. The current proportion was based on a shore A hardness of 50, which is an exploratory selection due to the lack of study on the shore hardness of bladder and rectum. Nonetheless, the range of prostate deformation magnitude, caused by the bladder and rectum deformation, covered the mean prostate motion during a radiotherapy course (0.18–0.59 cm) as reported in past work (Crook *et al* 1995, Britton *et al* 2005). On this account, we hold the view that having bladder and rectum models with similar mechanical properties to real ones would be a decided plus, but not a compulsory criterion for simulation of a realistic deformation scenario.

Through our study, it was found that the mean TREs increased with marker movements. In other words, the larger the deformations, the harder it is for the DIR algorithm to map each voxel accurately. Overall, DIR applying PBR as CROI achieved the most accurate voxel mapping, where the mean TREs were reduced to about half of the corresponding mean marker movements. The maximum mean TRE of  $0.309 \pm 0.038$  cm in the last deformation scenario with a mean marker movement of  $0.641 \pm 0.047$  cm is close to the TRE tolerance recommended by the American Association of Physicists in Medicine (AAPM) Task Group 132 (Brock *et al* 2017). This implies that with proper parameter tuning, the hybrid intensity and structure based deformable registration is capable of mapping dose with an accuracy that is adequate for clinical use.

The remarkably strong correlation between DSC and mean TRE for prostate (as presented in table 5) suggests that the contour mapping accuracy, reflected by DSC, serves as a reliable indicator of the voxel mapping accuracy within the homogenous medium. This could be attributed to the constant prostate volume in our phantom, providing a mass- and density-conserving deformation scenario. Under such condition, it has been shown that an accurate DIR-based dose mapping was achievable even in low contrast regions (Yeo *et al* 2012).

It is noteworthy that the correlation between DSC and mean TRE for bladder was relatively weaker compared to prostate and rectum. In figure 4, a cluster of blue points (representing bladder) were observed at  $DSC > 0.95$ , but the mean TREs did not remain consistently low and ranged up to 0.4 cm. This is

presumably due to the large volume of bladder, which allows a great degree of overlap between two contours, thereby easily yielding a high DSC. In fact, the volume dependence of DSC has been reported in a few studies (Deeley *et al* 2011, Kumarasiri *et al* 2014). Therefore, caution must be practiced when evaluating DIR accuracy for larger organs, especially when voxel mapping accuracy is of concern. In that case, contour-based metrics alone do not suffice for a conclusive QA result; additional metrics, such as TRE, should also be considered whenever possible.

Despite the past studies (Rohlfing 2012, Kirby *et al* 2013) which revealed the unreliability of contour-based metrics for assessing DIR accuracy, the results of our study appear to contradict this finding. One reason that could have contributed to this difference is the algorithms tested in the studies. Rohlfing designed a highly inaccurate algorithm which only excels in terms of certain surrogate measures of registration performance (Rohlfing 2012), while Kirby *et al* modified the algorithm smoothness factor using a beta version of MIM software (2013). Through their study, Kirby *et al* have emphasized the effect of regularization, which ensures a physically plausible deformation, on DIR accuracy (2013). This factor has also been accounted for in the RayStation algorithm, which objective function consists of an image similarity term, grid regularization terms, and anatomical penalty terms. In our study, the default parameters were used (other than specifying the CROI or FROI). Thus, our results suggest that the hybrid intensity and structure based deformable registration has an inherently good balance between the image information, physical plausibility of deformation, and the anatomical information (if provided by the user). In the case that an algorithm exhibits high inconsistency between contour-based metrics and voxel mapping accuracy, it may be worth tuning the regularization factor or introducing one, if it has not already been integrated into the algorithm.

As shown in figure 4, an upper limit of TRE of 0.3 cm, as recommended by the AAPM Task Group 132 (Brock *et al* 2017), corresponds to DSC > 0.85/0.90 for rectum/prostate. This suggested that for DIR QA where use of landmarks is not possible, a DSC greater than 0.85/0.90 should indicate a satisfactory voxel mapping accuracy for rectum/prostate (with the exception of mass- and density-changing scenario of a homogeneous medium, as discussed earlier). It should be noted, however, that this result is only applicable for the pelvis, and the hybrid intensity and structure based deformable registration on RayStation. Further research across diverse sites and algorithms is warranted to enhance the generalizability of these findings.

One main limitation of our study is the overriding of CT number in the phantom images. While it is an indispensable step to improve their resemblance to real patient images, the overriding of the region surrounding the organs to the adipose mean CT number has itself posed drawback in the analysis. This is because the contrast between the organs and the surrounding medium has increased beyond realistic cases where the organs (especially prostate) could be less distinguishable. Nevertheless, the low mean TRE of the markers inside prostate as mentioned earlier could attest to the ability of the algorithm to register accurately, not just in high-contrast regions, but also in low-contrast regions.

Depending on the dose distribution in a certain region, the effect of voxel mapping error on the dose mapping error would also vary. For instance, in a region with high dose gradients, a slight voxel mapping error would give rise to a huge discrepancy in the mapped dose distribution. This relationship could be investigated by simulating a radiotherapy treatment course where organ deformations take place, and subsequently checking the dose discrepancy between the measured and mapped dose distributions. Since the irradiation of the Polyjet digital material with doses up to 100 Gy did not alter its mechanical properties (indicating a negligible radiation damage), this deformable phantom (with the markers replaced by tiny dosimeters) could be adaptable for the aforementioned study. This is one of the future work that we are looking to explore in the future.

## 5. Conclusion

In this work, we demonstrated the possibility of building a deformable anthropomorphic pelvis phantom using 3D printing materials. Through simulation of various deformation scenarios, it has been shown that the DIR voxel mapping accuracy decreases with increasing deformation magnitude. DSC was found to be a good indicator of voxel mapping accuracy for prostate and rectum, but a comparatively poorer one for bladder. Hence, additional metrics should be taken into account for evaluation of bladder to ensure a reliable QA result. As different DIR algorithms make use of different optimization parameters, similar analyses can be performed for other algorithms to cross-check these findings.

## Data availability statement

All data that support the findings of this study are included within the article (and any supplementary information files).

## Conflict of interest

The authors declare no direct conflict of interests.

## Funding support

Hong Qi Tan is supported by the Duke-NUS Oncology Academic Program Goh Foundation Proton Research Programme (08/FY2023/EX(SL)/163-A218(b)), Clinical & Systems Innovation Support—Innovation Seed Grant (08/FY2022/P2/02-A68)

## Author contribution statement

Study conception and design: Yun Ming Wong, Hong Qi Tan

Data acquisition and analysis: Yun Ming Wong, Calvin Wei Yang Koh, Kah Seng Lew, Clifford Ghee Ann Chua, Ping Lin Yeap, Ee Teng Zhang, Ashley Li Kuan Ong, Jeffrey Kit Loong Tuan, Hong Qi Tan

Data interpretation: All authors

Statistical analyses: Yun Ming Wong, Hong Qi Tan

Obtained funding: Hong Qi Tan

Administrative, technical, or material support: Hong Qi Tan, Calvin Wei Yang Koh, Ee Teng Zhang

Study supervision: Bing Feng Ng, Wen Siang Lew, James Cheow Lei Lee

Drafting of manuscript: Yun Ming Wong

Approval of final manuscript: All authors

## ORCID iDs

Yun Ming Wong  <https://orcid.org/0000-0002-1564-2087>

Hong Qi Tan  <https://orcid.org/0000-0001-7878-4544>

## References

- 3D slicer image computing platform n.d. 3D Slicer (available at: <https://slicer.org/>) (Accessed 1 November 2024)
- Agarwal J P, Tibdewal A, Johnny C, Mummudi N, Prabhaskar K, Sahoo P and Kinshikar R 2020 Adaptive radiotherapy in locally advanced lung cancers—real world scenario from a tertiary cancer center *Int. J. Radiat. Oncol. Biol. Phys.* **108** e128–9
- Avkshtol V, Meng B, Shen C, Choi B S, Okoroafor C, Moon D, Sher D and Lin M-H 2023 Early experience of online adaptive radiation therapy for definitive radiation of patients with head and neck cancer *Adv. Radiat. Oncol.* **8** 101256
- Bohoudi O, Lagerwaard F J, Bruynzeel A M E, Niebuhr N I, Johnen W, Senan S, Slotman B J, Pfaffenberger A and Palacios M A 2019 End-to-end empirical validation of dose accumulation in MRI-guided adaptive radiotherapy for prostate cancer using an anthropomorphic deformable pelvis phantom *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **141** 200–7
- Britton K R, Takai Y, Mitsuya M, Nemoto K, Ogawa Y and Yamada S 2005 Evaluation of inter- and intrafraction organ motion during intensity modulated radiation therapy (IMRT) for localized prostate cancer measured by a newly developed on-board image-guided system *Radiat. Med.* **23** 14–24
- Brock K K, Mutic S, McNutt T R, Li H and Kessler M L 2017 Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy committee task group no. 132 *Med. Phys.* **44** e43–e76
- Buciuman N and Marcu L G 2022 Adaptive radiotherapy in head and neck cancer using volumetric modulated arc therapy *J. Pers. Med.* **12** 668
- Cao Y, Chen H, Gogineni E, Li H and Deville C 2022 Initial experience with real-time gated proton therapy (RGPT) in the definitive treatment of prostate cancer *Int. J. Radiat. Oncol.* **114** e542–3
- Castillo R, Castillo E, Guerra R, Johnson V E, McPhail T, Garg A K and Guerrero T 2009 A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets *Phys. Med. Biol.* **54** 1849–70
- Chang C-W, Nilsson R, Andersson S, Bohannon D, Patel S A, Patel P R, Liu T, Yang X and Zhou J 2023 An optimized framework for cone-beam computed tomography-based online evaluation for proton therapy *Med. Phys.* **50** 5375–86
- Christensen M B, Oberg K and Wolchok J C 2015 Tensile properties of the rectal and sigmoid colon: a comparative analysis of human and porcine tissue *SpringerPlus* **4** 142
- Christiansen R L et al 2022 Online adaptive radiotherapy potentially reduces toxicity for high-risk prostate cancer treatment *Radiother. Oncol.* **167** 165–71
- Crook J M, Raymond Y, Salhani D, Yang H and Esche B 1995 Prostate motion during standard radiotherapy as assessed by fiducial markers *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **37** 35–42
- Dahms S E, Piechota H J, Dahiya R, Lue T F and Tanagho E A 1998 Composition and biomechanical properties of the bladder acellular matrix graft: comparative analysis in rat, pig and human *Br. J. Urol.* **82** 411–9
- Dawson L A, Mah K, Franssen E and Morton G 1998 Target position variability throughout prostate radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **42** 1155–61
- Deeley M A et al 2011 Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study *Phys. Med. Biol.* **56** 4557–77
- Hoppen L, Sarria G R, Kwok C S, Boda-Heggemann J, Buergy D, Ehmann M, Giordano F A and Fleckenstein J 2023 Dosimetric benefits of adaptive radiation therapy for patients with stage III non-small cell lung cancer *Radiat. Oncol.* **18** 34
- Irmak S, Georg D and Lechner W 2020 Comparison of CBCT conversion methods for dose calculation in the head and neck region *Z. Für Med. Phys.* **30** 289–99

- Kirby N, Chuang C, Ueda U and Pouliot J 2013 The need for application-based adaptation of deformable image registration *Med. Phys.* **40** 011702
- Kotte A N T J, Hofman P, Lagendijk J J W, van Vulpen M and van der Heide U A 2007 Intrafraction motion of the prostate during external-beam radiation therapy: analysis of 427 patients with implanted fiducial markers *Int. J. Radiat. Oncol. Biol. Phys.* **69** 419–25
- Kumarasiri A, Siddiqui F, Liu C, Yechieli R, Shah M, Pradhan D, Zhong H, Chetty I J and Kim J 2014 Deformable image registration based automatic CT-to-CT contour propagation for head and neck adaptive radiotherapy in the routine clinical setting *Med. Phys.* **41** 121712
- Kupelian P, Willoughby T, Litzenberg D, Sandler H, Roach M, Levine L, van Waardenburg M, Cunningham A and Meeks S 2005 Clinical experience with the Calypso® 4D localization system in prostate cancer patients: implantation, tolerance, migration, localization and real time tracking *Int. J. Radiat. Oncol. Biol. Phys.* **63** S197
- Litzenberg D W, Balter J M, Hadley S W, Sandler H M, Willoughby T R, Kupelian P A and Levine L 2006 Influence of intrafraction motion on margins for prostate radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **65** 548–53
- Møller D S et al 2022 Survival benefits for non-small cell lung cancer patients treated with adaptive radiotherapy *Radiother. Oncol.* **168** 234–40
- Nederveen A J, van der Heide U A, Dehnad H, van Moorselaar R J A, Hofman P and Lagendijk J J W 2002 Measurements and clinical consequences of prostate motion during a radiotherapy fraction *Int. J. Radiat. Oncol. Biol. Phys.* **53** 206–14
- Niebuhr N I, Johnen W, Echner G, Runz A, Bach M, Stoll M, Giske K, Greilich S and Pfaffenberger A 2019 The ADAM-pelvis phantom—an anthropomorphic, deformable and multimodal phantom for MRgRT *Phys. Med. Biol.* **64** 04NT05
- Nijkamp J, Pos F J, Nuver T T, de Jong R, Remeijer P, Sonke J-J and Lebesque J V 2008 Adaptive radiotherapy for prostate cancer using kilovoltage cone-beam computed tomography: first clinical results *Int. J. Radiat. Oncol. Biol. Phys.* **70** 75–82
- Paganelli C, Meschini G, Molinelli S, Riboldi M and Baroni G 2018 Patient-specific validation of deformable image registration in radiation therapy: overview and caveats *Med. Phys.* **45** e908–22
- Pang E P P, Knight K, Fan Q, Tan S X F, Ang K W, Master Z, Mui W-H, Leung R W-K, Baird M and Tuan J K L 2018 Analysis of intra-fraction prostate motion and derivation of duration-dependent margins for radiotherapy using real-time 4D ultrasound *Phys. Imaging Radiat. Oncol.* **5** 102–7
- Rigaud B, Simon A, Castelli J, Gobeli M, Ospina Arango J-D, Cazoulat G, Henry O, Haigron P and De Crevoisier R 2015 Evaluation of deformable image registration methods for dose monitoring in head and neck radiotherapy *BioMed Res. Int.* **2015** e726268
- Rigaud B, Simon A, Castelli J, Lafond C, Acosta O, Haigron P, Cazoulat G and de Crevoisier R 2019 Deformable image registration for radiation therapy: principle, methods, applications and evaluation *Acta Oncol.* **58** 1225–37
- Roeske J C, Forman J D, Mesina C F, He T, Pelizzari C A, Fontenla E, Vijayakumar S and Chen G T 1995 Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **33** 1321–9
- Rohlfing T 2012 Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable *IEEE Trans. Med. Imaging* **31** 153–63
- Schwartz D L et al 2013 Adaptive radiotherapy for head and neck cancer—dosimetric results from a prospective clinical trial *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **106** 80–84
- Sheikh K, Liu D, Li H, Acharya S, Ladra M M and Hrinivich W T 2022 Dosimetric evaluation of cone-beam CT-based synthetic CTs in pediatric patients undergoing intensity-modulated proton therapy *J. Appl. Clin. Med. Phys.* **23** e13604
- Thing R S, Nilsson R, Andersson S, Berg M and Lund M D 2022 Evaluation of CBCT based dose calculation in the thorax and pelvis using two generic algorithms *Phys. Med.* **103** 157–65
- Varadhan R, Karangelis G, Krishnan K and Hui S 2013 A framework for deformable image registration validation in radiotherapy clinical applications *J. Appl. Clin. Med. Phys.* **14** 192–213
- Varadhan R, Magome T and Hui S 2016 Characterization of deformation and physical force in uniform low contrast anatomy and its impact on accuracy of deformable image registration *Med. Phys.* **43** 52–61
- Wang H, Dong L, Lii M F, Lee A L, de Crevoisier R, Mohan R, Cox J D, Kuban D A and Cheung R 2005 Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **61** 725–35
- Waters M et al 2024 CT-based online adaptive radiotherapy improves target coverage and organ at risk (OAR) avoidance in stereotactic body radiation therapy (SBRT) for prostate cancer *Clin. Transl. Radiat. Oncol.* **44** 100693
- Yan D, Vicini F, Wong J and Martinez A 1997 Adaptive radiation therapy *Phys. Med. Biol.* **42** 123–32
- Yeo U J, Taylor M L, Supple J R, Smith R L, Dunn L, Kron T and Franich R D 2012 Is it sensible to “deform” dose? 3D experimental validation of dose-warping *Med. Phys.* **39** 5065–72
- Zhang Z 2023 Liver vessel segmentation and accurate landmark pairs detection for quantitative liver deformable image registration verification