Evaluation of Deformable Image Registration-Based Contour Propagation From Planning CT to Cone-Beam CT

Andrew J. Woerner, MS[1], Mehee Choi, MD[1], Matthew M. Harkenrider, MD[1],
John C. Roeske, PhD[1], and Murat Surucu, PhD[1]

Abstract

Purpose: We evaluated the performance of organ contour propagation from a planning computed tomography to cone-beam computed tomography with deformable image registration by comparing contours to manual contouring. **Materials and Methods:** Sixteen patients were retrospectively identified based on showing considerable physical change throughout the course of treatment. Multiple organs in the 3 regions (head and neck, prostate, and pancreas) were evaluated. A cone-beam computed tomography from the end of treatment was registered to the planning computed tomography using rigid registration, followed by deformable image registration. The contours were copied on cone-beam computed tomography image sets using rigid registration and modified by 2 radiation oncologists. Contours were compared using Dice similarity coefficient, mean surface distance, and Hausdorff distance. **Results:** The mean physician-to-physician Dice similarity coefficient for all organs was 0.90. When compared to each physician's contours, the overall mean for rigid was 0.76 ($P < .001$), and it was improved to 0.79 ($P < .001$) for deformable image registration. Comparing deformable image registration to physicians resulted in a mean Dice similarity coefficient of 0.77, 0.74, and 0.84 for head and neck, prostate, and pancreas groups, respectively; whereas, the physician-to-physician mean agreement for these sites was 0.87, 0.90, and 0.93 ($P < .001$, for all sites). The mean surface distance for physician-to-physician contours was 1.01 mm, compared to 2.58 mm for rigid-to-physician contours and 2.24 mm for deformable image registration-to-physician contours. The mean physician-to-physician Hausdorff distance was 11.32 mm, and when compared to any physician's contours, the mean for rigid and deformable image registration was 12.1 mm and 12.0 mm ($P < .001$), respectively. **Conclusion:** The physicians had a high level of agreement via the 3 metrics; however, deformable image registration fell short of this level of agreement. The automatic workflows using deformable image registration to deform contours to cone-beam computed tomography to evaluate the changes during treatment should be used with caution.

Keywords

deformable image registration, cone-beam CT, adaptive radiotherapy, contour deformation, deformation error

Abbreviations

ART, adaptive radiotherapy; CBCT, cone-beam computed tomography; CT, computed tomography; DIR, deformable image registration; DSC, Dice similarity coefficient; H&N, head and neck; HD, Hausdorff distance; MSD, mean surface distance; OARs, organs at risk; pCT, planning CT; RO, radiation oncologist

Introduction

Changes in patient anatomy can alter the complex radiotherapy treatment plans that are based on the planning computed tomography (pCT) scan performed before the start of treatment. These changes may lead to altered dose–volume parameters that prevent delivering the intended doses to targets and increase the toxic dose delivered to the organs at risk

[1] Department of Radiation Oncology, Loyola University Medical Center, Maywood, IL, USA

Corresponding Author:
Murat Surucu, PhD, Department of Radiation Oncology, Loyola University Medical Center, Rm 2944, 2160 S 1st Ave, Maguire Center, Maywood, IL 60153, USA.
Email: msurucu@lumc.edu

(OARs).[1-4] Radiotherapy plans can be adapted to the altered anatomy in order to safely deliver the prescription dose to the tumor, while ensuring OAR sparing. This adaptive radiotherapy (ART) requires continual monitoring of the target and OAR regions for evaluation and replanning. However, manual recontouring is highly time consuming, with some estimates close to 3 hours to recontour a single head and neck (H&N) case.[5]

The use of online cone-beam computed tomography (CBCT) on a daily basis during treatment has shown great success at revealing changes in patient anatomy relative to the pCT.[6,7] To enhance the efficiency of treatment delivery, computed propagation of the target and OAR contours between 2 image sets can be performed using image registration.[8] Overall, the addition of CBCT to ART has contributed to a reduction in the volume of healthy tissue irradiated and a reduction in the dose delivered to OAR, because it may alert physicians to consider adapting the plan to the altered anatomy.[9,10]

Two forms of image registration exist—rigid registration and deformable image registration (DIR). Rigid registration is limited to 6 degrees of freedom via translations and rotations, so these may not produce acceptable results if there are non-rigid tissue deformations between 2 image sets. On the other hand, DIR can allow for elastic and nonrigid deformations. These deformations are capable of locally warping the target image to align with the reference image. Because of this, DIR performs more accurate image modifications than rigid registration methods and is becoming a growing tool in ART.[11-13] Studies continue to show increasing accuracy of the multitude of available DIR algorithms and promise for increased use with automatic target delineation on CBCT images.[3,14]

With the expansion of available software using DIR algorithms, so too has there been an increase in the study of qualitative evaluation of these algorithms in the clinical setting. Firstly, DIR has shown a great deal of variation of uncertainties due to the choice of algorithm used, as well as varying degrees of challenges to image quality.[15,16] Also, DIR still falls considerably short of manual recontouring in terms of agreement.[13] In a recent study by Perna et al, the use of DIR in the context of pelvic radiotherapy showed a low level of conformality compared to that of manually drawn contours, but a much higher level of agreeability when the DIR contours were manually adjusted. These manual adjustments resulted in a 50% reduction in the time spent on contouring images.[17] In the context of H&N cases, a study by Kumarasiri et al emphasized the characteristics of the OARs (volume and boundary definition) when it came to the level of conformality of DIR contours, with the highest conformality seen in larger organs with clear boundaries.[18]

The goal of this study is to assess DIR performance against manually drawn contours of radiation oncologists (ROs), in order to further speculate on DIR's expanding clinical applicability. Since physicians use multiple image modalities to contour target volumes, and these additional scans were not repeated for this study, we limited our analysis to OAR contours, due to their greater definition than target volumes.

Table 1. Demographic Information of Analyzed Patients and the Organs Identified for Registration.

Patient Set	Organs Included for Analysis
H&N	Esophagus
6 patients	Mandible
Mean time elapsed: 38.5 days	Left parotid
Median total dose: 49.3 Gy	Right parotid
	Spinal cord
Prostate	Bladder
5 patients	Prostate
Mean time elapsed: 46.2 days	Rectum
Median total dose: 78 Gy	Seminal vesicles
Pancreas	Liver
5 patients	Left kidney
Mean time elapsed: 19.6 days	Right kidney
Median total dose: 45 Gy	Spinal cord

Abbreviation: H&N, head and neck.

Unique to this study is the fact that 3 organ regions are included (H&N, prostate, and pancreas) in order to assess region-specific applicability, as well as to highlight and compare unique challenges encountered between regions. Also, inclusion of 3 metrics to assess the performance of image registrations can help address possible drawbacks to individual metric choice.

Materials and Methods

Patients were selected retrospectively with the approval of the Loyola University Medical Center's institutional review board. A total of 16 were selected (6 H&N, 5 prostate, and 5 pancreas) based upon showing physical change throughout the course of treatment. The treatment pCT was compared to a treatment CBCT that was taken near the end of treatment; mean amount of time elapsed between images was 35 days (8-66 days). For each set of patients, 4 or 5 organs of interest were selected to be registered. Organ at risk contours on the pCT were initially contoured by the department's dosimetrists at the time of treatment planning and then refined by a treating RO. Rigid registrations of these contours were applied to the follow-up CBCT and provided to the ROs for manual contouring, regardless of who the primary RO responsible for the original contours. The organ sets chosen for each region are listed Table 1.

Image registrations were performed using Velocity Advanced Imaging 2.8.1 (Varian Medical Systems, Palo Alto, California). Initially, a rigid registration was performed based on the region of interest's bony anatomy (eg, cervical spine in patients with H&N, pelvis and femoral heads in patients with prostate cancer, and thoracolumbar spine in patients with pancreatic cancer). Using the rigid registration as a reference, subsequent DIRs were performed with the software's modified B-spline deformable with mutual information-based matching algorithm. An initial "extended deformable multi-pass" (6-pass deformable registration with progressively finer resolution) deformation was performed using the entire CBCT

Table 2. Mean Values for the 3 Metrics Across All Patients and All Organs Contoured.

Metric	Rigid Versus RO1	Rigid Versus RO2	DIR Versus RO1	DIR Versus RO2	RO1 Versus RO2
DSC	0.75 ($\pm$0.16)	0.75 ($\pm$0.16)	0.78 ($\pm$0.13)	0.78 ($\pm$0.14)	0.90 ($\pm$0.09)
MSD (mm)	2.51 ($\pm$1.80)	2.64 ($\pm$1.90)	2.20 ($\pm$1.31)	2.28 ($\pm$1.40)	1.01 ($\pm$0.82)
HD (mm)	11.84 ($\pm$6.16)	14.27 ($\pm$12.87)	12.35 ($\pm$6.79)	14.26 ($\pm$12.27)	11.32 ($\pm$13.00)

Abbreviations: DIR, deformable image registration; DSC, Dice similarity coefficient; HD, Hausdorff distance; MSD, mean surface distance; RO1, radiation oncologist 1; RO2, radiation oncologist 2.

Table 3. Mean Values for the 3 Metrics for Physician-to-Physician Comparisons.

Metric	Overall	H&N	Prostate	Pancreas
DSC	0.90 ($\pm$0.09)	0.87 ($\pm$0.11)	0.90 ($\pm$0.06)	0.93 ($\pm$0.07)
MSD (mm)	1.01 ($\pm$0.82)	0.88 ($\pm$0.83)	0.99 ($\pm$0.44)	1.21 ($\pm$1.05)
HD (mm)	11.32 ($\pm$13.00)	7.93 ($\pm$6.21)	8.12 ($\pm$4.52)	18.92 ($\pm$20.47)

Abbreviations: DSC, Dice similarity coefficient; H&N, head and neck; HD, Hausdorff distance; MSD, mean surface distance.
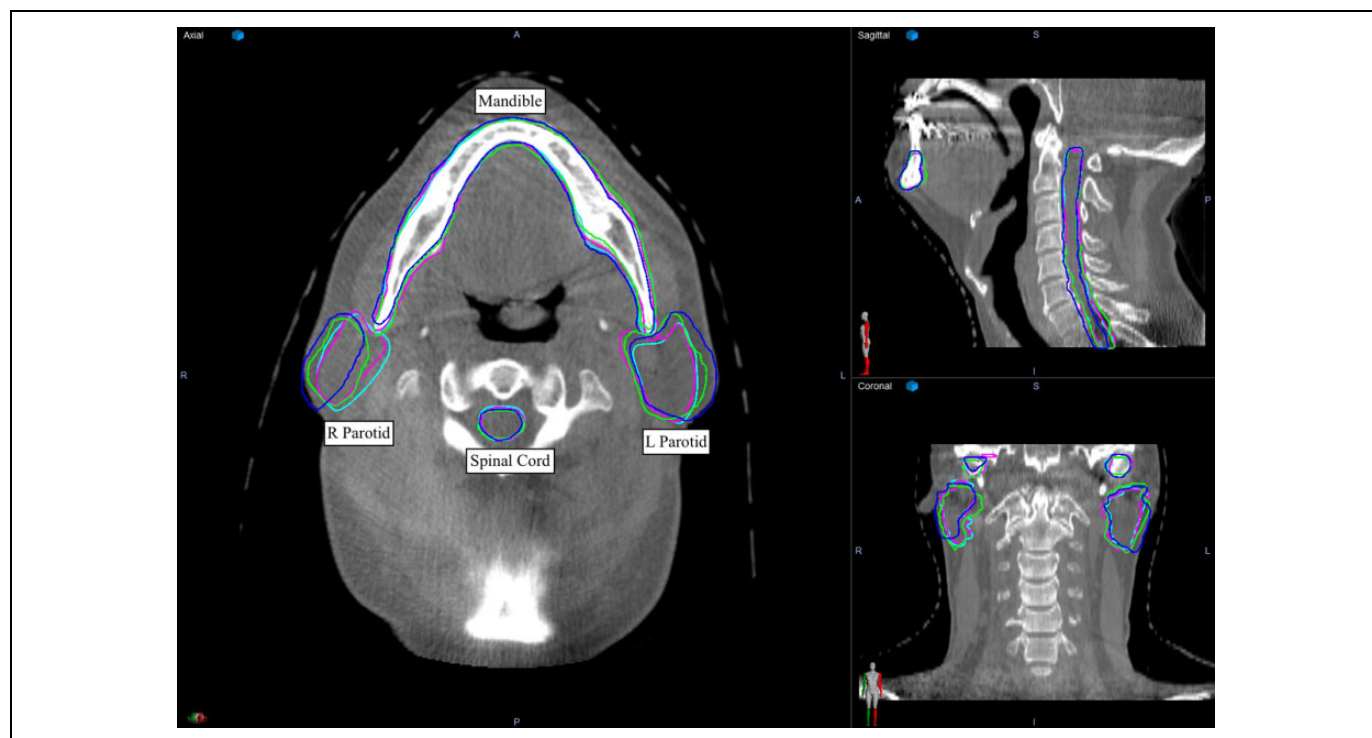

Figure 1. Representative OAR contours of the H&N region, contoured on mid-treatment CBCT. This highlights the 4 OAR contour sets of comparison in axial, sagittal, and coronal views. CBCT indicates cone-beam computed tomography; dark blue, rigid registration; green, deformable image registration; H&N, head and neck; light blue, RO1; magenta, RO2; OAR, organ at risk.

image as the region of interest. All DIR image sets were closely visually inspected by 2 of the authors to look for any readily apparent aberrancies. Then, if needed, an additional single-pass deformable registration set to "fine" resolution was performed, using a narrower region of interest that included the organs of interest.

To assess the accuracy of the registered organ contours, 2 ROs manually adjusted the organs of interest from contours that were rigidly registered to the CBCT image sets. Organ contours were compared using 3 metrics—Dice similarity coefficient (DSC), mean surface distance (MSD), and Hausdorff distance (HD). The DSC measures the overlap and volume shared by 2 structures. A result of 0 means no overlap, and a result of 1 means perfect overlap.[19] In order to compute MSD, the smallest distance from 1 point on a contour X to all points of contour Y is calculated, and MSD becomes the average of
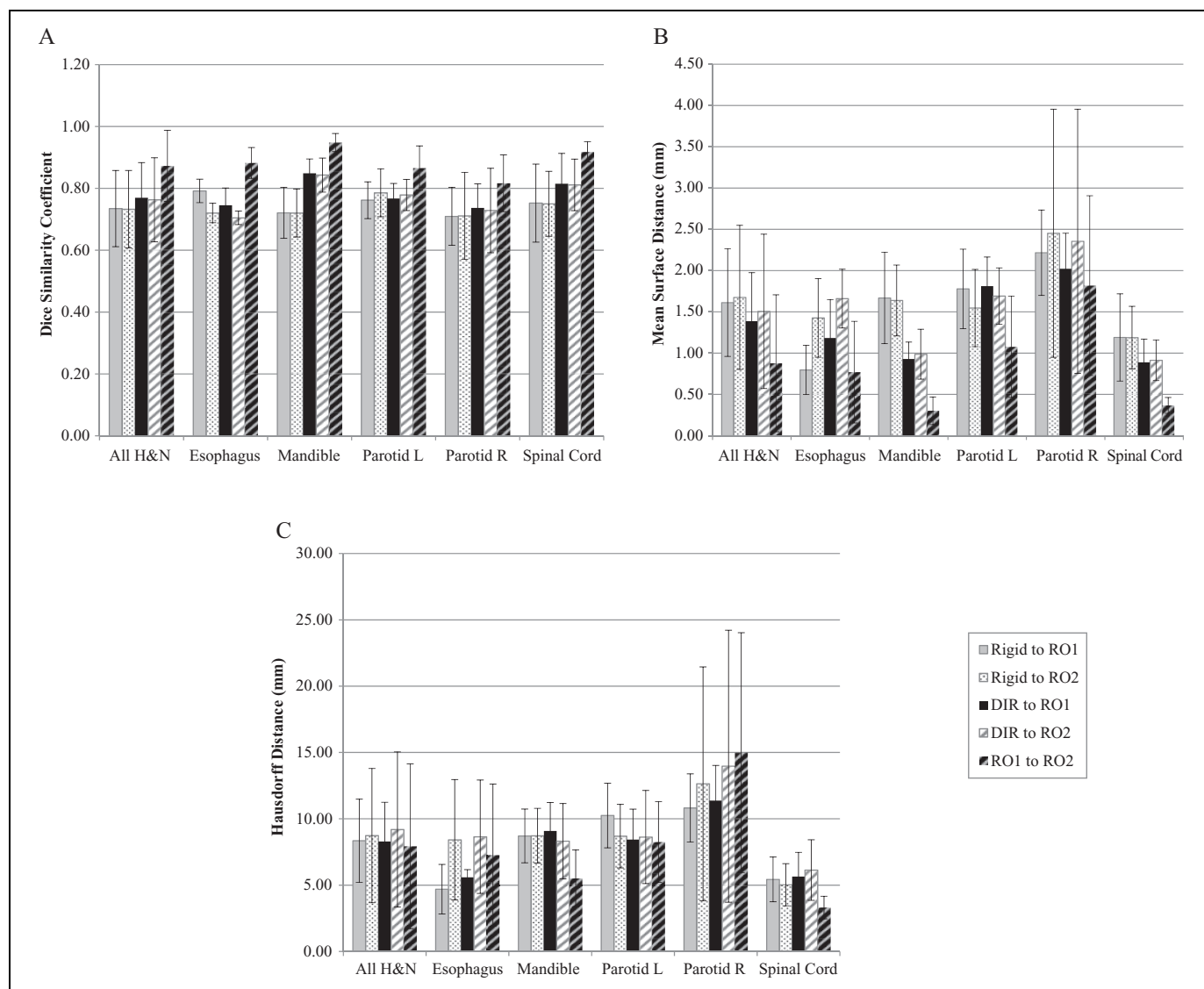
Figure 2. The DIR performance in the context of H&N cancer. A, Mean DSC values for the H&N region across the contoured organs. B, Mean MSD for the H&N region. C, Mean HD values for the H&N region. DIR indicates deformable image registration; DSC, Dice similarity coefficient; H&N, head and neck; HD, Hausdorff distance; MSD, mean surface distance; RO1, radiation oncologist 1; RO2, radiation oncologist 2.

the smallest distances for all points on contour X. The HD is a measure of the maximum distance between 2 objects.[20] Student *t* test with an α level of .05 was used to calculate statistical significance.

Results

Physician-to-Physician Comparison

For all organs across the 3 regions, the mean DSC, MSD, and HD were 0.90 ($\pm$0.09), 1.01 mm ($\pm$0.82), and 11.32 mm ($\pm$13.0), respectively (Table 2), when looking at physician (RO1) to physician (RO2) comparisons.

The region showing the highest DSC between physicians was the pancreas, with a mean DSC of 0.93 (Table 3). On the other hand, the region showing the lowest DSC was the H&N

with a mean DSC of 0.87 (the prostate region showed a mean DSC of 0.90). This trend was reversed for both the MSD and HD metrics, with the H&N region having superior numbers versus the pancreas region showing the worst agreement. As far as particular organ contours, the mandible and liver both showed the highest DSC, with a mean of 0.95; the seminal vesicles showed the lowest DSC agreement with a mean of 0.84.

Although the H&N region had the lowest overall mean MSD, it also included the organ with the greatest MSD, the right parotid, 1.82 ($\pm$1.09 mm). However, the mandible showed the lowest MSD, 0.30 ($\pm$0.17 mm), of all organ contours. Overall MSD values are shown in Table 3. As for organ specifics of the HD metric, the spinal cord (as part of the H&N region) had the lowest mean HD (smaller HD is ideal) of 3.31 mm and the liver had the greatest mean HD of 26.17 mm.
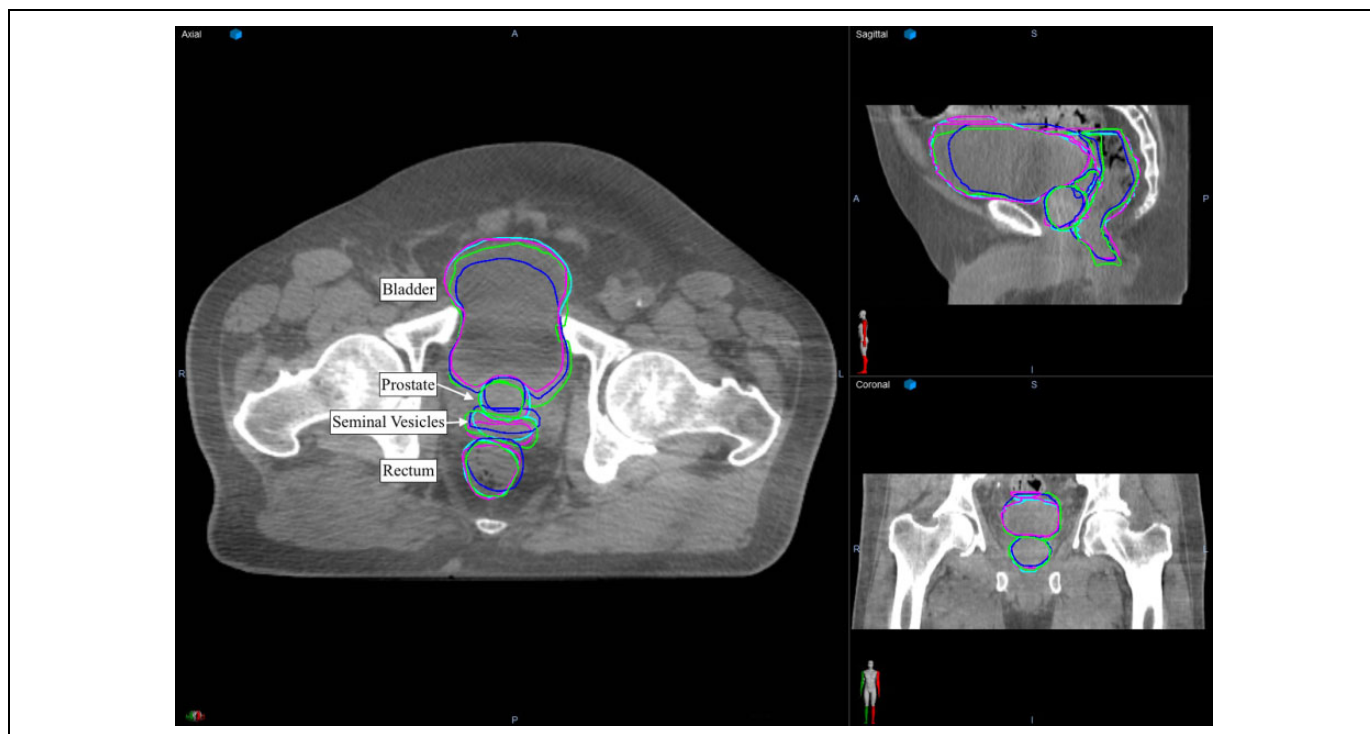
Figure 3. Representative OAR contours of the prostate region, contoured on mid-treatment CBCT. This highlights the 4 OAR contour sets of comparison in axial, sagittal, and coronal views. CBCT indicates cone-beam computed tomography; dark blue, rigid registration; green, deformable image registration; light blue, RO1; magenta, RO2; OAR, organ at risk.

Head and Neck

There were 6 patient image sets included for the H&N region. For the contours of this region, the use of DIR produced a mean improvement from rigid registration by 0.04 mm ($\pm$0.08 mm) and 0.20 mm ($\pm$0.52 mm) for the DSC and MSD metrics, respectively. A representative H&N case demonstrating the rigid and deformable image-registered contours as well as the ones modified by 2 ROs is shown in Figure 1. The use of DIR lead to a mean increase of 0.21 mm ($\pm$2.21 mm) for the HD metric. Of the 3 organ regions, the H&N contour comparisons of DIR to the RO1 or RO2 conformality was closest to the RO1 to RO2 conformality when measured by the MSD and HD metrics (within 0.63 mm and 1.26 mm, respectively).

In terms of the DSC measurements, as previously mentioned, the H&N region posted the lowest DSC conformality of RO1 to RO2 contours. The H&N region's mean DSC values for DIR to RO1 or RO2 were 0.77 ($\pm$0.11) and 0.76 ($\pm$0.14), respectively. These values were improvements from rigid registration by 0.04 and 0.03 but less than RO1 to RO2 agreeability by 0.10 and 0.11. The esophagus proved to be the most challenging for the DIR contours reporting DSC values of 0.75 ($\pm$0.06) and 0.70 ($\pm$0.02), compared to RO1 to RO2 DSC of 0.88 ($\pm$0.05). On the other hand, the mandible had the highest DSC values of the region with 0.85 ($\pm$0.05) and 0.84 ($\pm$0.06) for DIR.

The H&N region had the lowest overall MSD values of the 3 regions included. For DIR to RO1 or RO2 analysis, mean MSD values were 1.39 mm ($\pm$0.59 mm) and 1.51 mm ($\pm$0.93 mm).

These values were improvements from rigid registration by 0.22 and 0.17 mm but greater than RO1 to RO2 agreeability by 0.51 and 0.63 mm. Much like the DSC metric, the mandible had some of the better MSD values of the region, but here the spinal cord had the lowest mean MSD values for DIR to RO1 or RO2 conformality, 0.89 mm ($\pm$0.28 mm) and 0.91 mm ($\pm$0.25 mm).

As for the HD measurements, the H&N region once again had the lowest overall values of all 3 regions. The mean HD of DIR to RO1 or RO2 was 8.29 mm ($\pm$2.95 mm) and 9.19 mm ($\pm$5.85 mm), compared to the RO1 to RO2 mean HD of 7.93 mm ($\pm$6.21 mm). For this metric, the right parotid posted the worst values of all comparisons, and the lowest HDs were seen with the spinal cord. For detailed information about the contours in this region, refer to Figure 2.

Prostate

The prostate region included contours from 5 patients. Overall, this region showed the worst levels of agreement of DIR to physician-drawn contours. A representative prostate case is shown in Figure 3, where the bladder filling on the treatment day being different than the day of pCT is seen, by comparing the rigid registered (blue) bladder contour. Notably, the seminal vesicles proved to be a challenge for the registration software, posting the worst values of all 3 metrics across contours from all regions when comparing registered images to physician-contoured images. The physician-to-physician
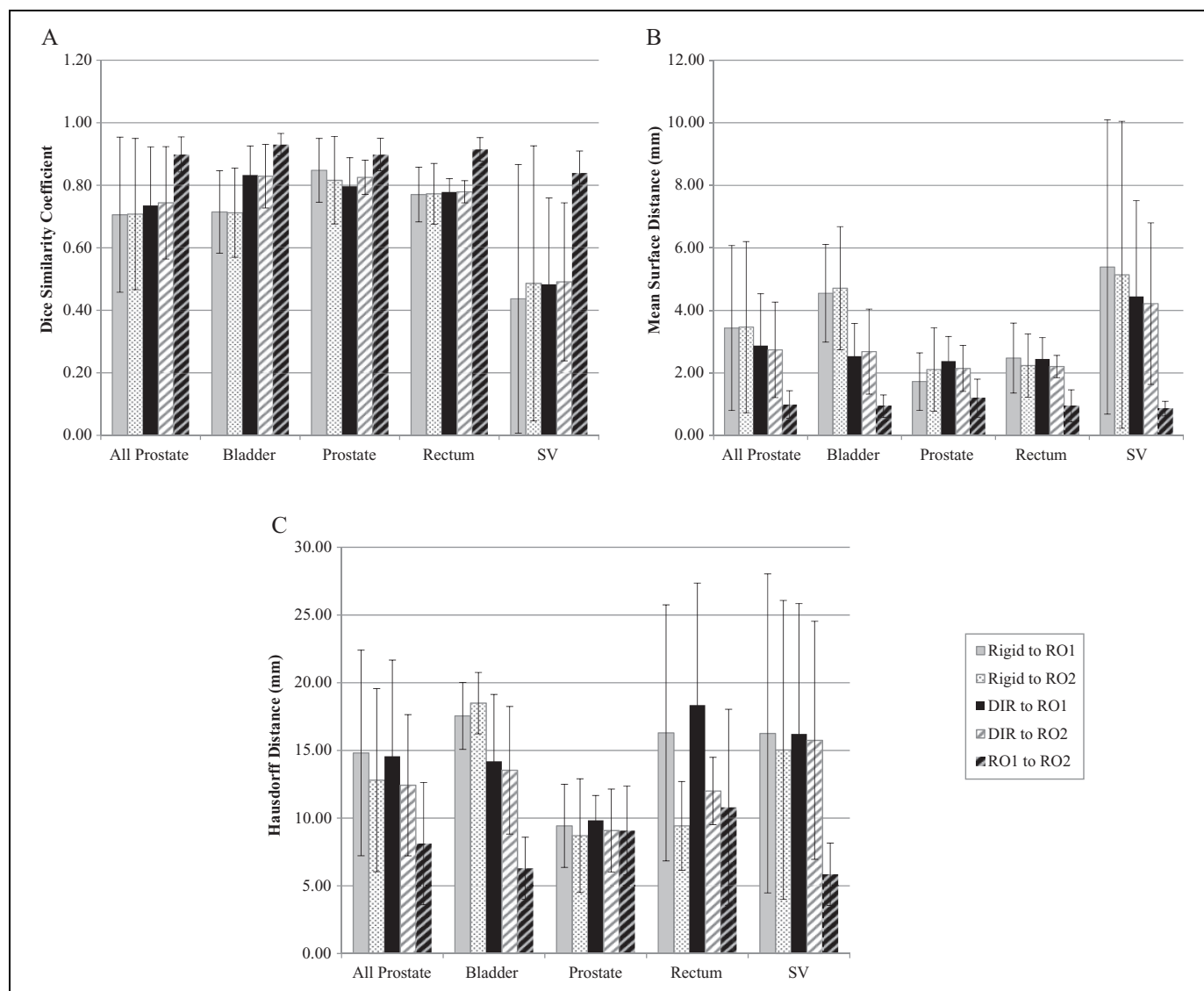
Figure 4. The DIR performance in the context of prostate cancer. A, Mean DSC values of the prostate region across the contoured organs. B, Mean MSD values of the prostate region. C, Mean HD values of the prostate region. DIR indicates deformable image registration; DSC, Dice similarity coefficient; HD, Hausdorff distance; MSD, mean surface distance; RO1, radiation oncologist 1; RO2, radiation oncologist 2.

agreement for the seminal vesicles showed the second lowest mean DSC, 0.84 ($\pm$0.07) of all organ contours (second only to another soft tissue organ, the right parotid). On the other hand, the bladder contours showed one of the highest improvements in mean DSC values when comparing DIR to rigid registration (+0.12 improvement).

When looking at mean MSD values for the region, DIR showed improvement compared to rigid registration for all organ contours except for the prostate. The mean improvement from DIR shown with the MSD metric was 0.56 and 0.73 mm for each physician's contours. However, the DIR to RO1 or RO2 MSD values were still 1.89 and 1.75 mm greater than the RO1 to RO2 MSD value, 0.99 mm ($\pm$0.44 mm).

Analysis of the prostate region using the HD metric continued to show modest improvement with the use of DIR over rigid registrations. These improvements were 0.24 and 0.38

mm for RO1 and RO2, respectively. However, the DIR-to-physician comparisons showed HD values that were 6.45 and 4.31 mm greater than the RO1 to RO2 comparison. For further detail in regard to the 3 metrics for the region, refer to Figure 4.

Pancreas

The pancreas region included analysis of 5 patients. This region showed the highest DSC values for all comparisons within the region, and the DIR to RO1 or RO2, 0.84 ($\pm$0.06) and 0.84 ($\pm$0.05), was the closest to the RO1 to RO2 DSC, 0.93 ($\pm$0.07; representative contours are shown in Figure 5). However, the use of DIR only produced a DSC improvement of 0.01 when compared to rigid registration. Of all the DSC values, organ contours from this region (left kidney, right kidney, and the liver) all had greater agreeability of DIR to RO1 or RO2
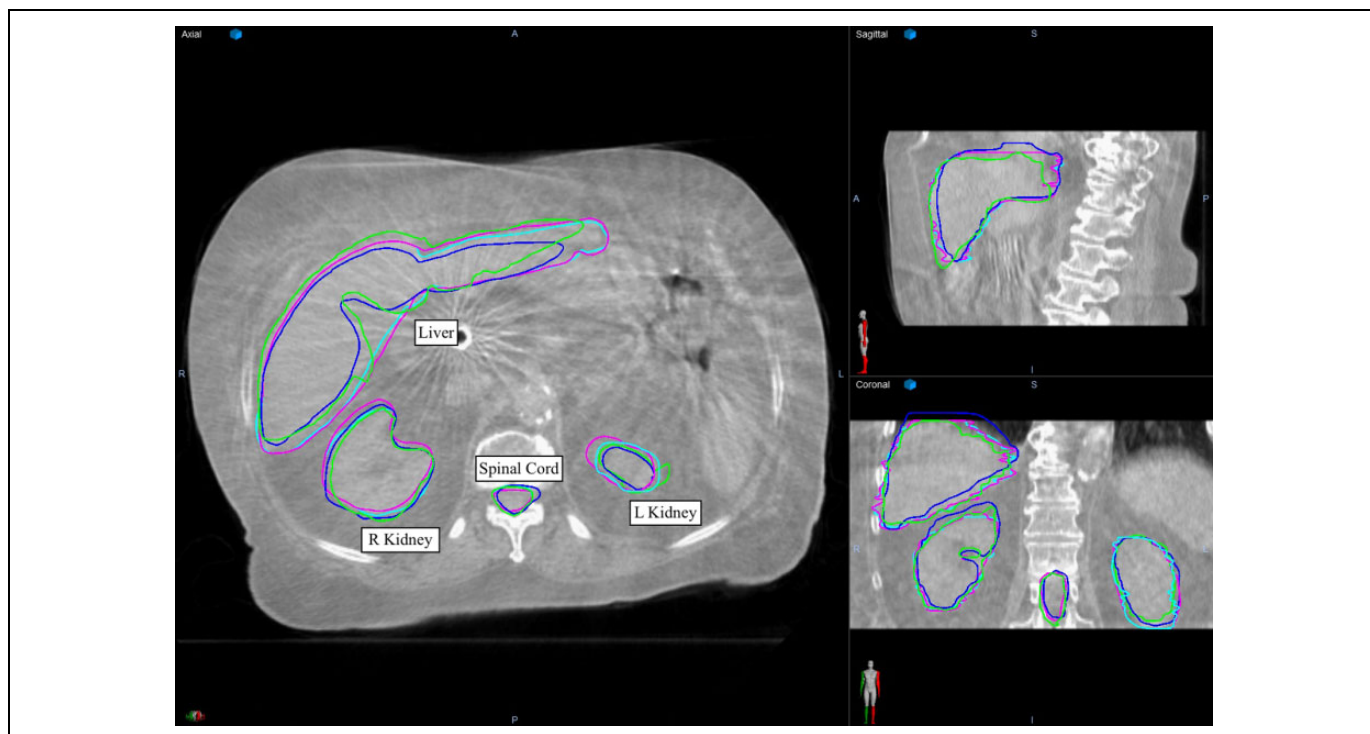
Figure 5. Representative OAR contours of the pancreas region, contoured on mid-treatment CBCT. This highlights the 4 OAR contour sets of comparison in axial, sagittal, and coronal views. CBCT indicates cone-beam computed tomography; dark blue, rigid registration; green, deformable image registration; light blue, RO1; magenta, RO2; OAR, organ at risk.

using the DSC metric than any of the other organs across the 3 regions.

The use of the MSD metric displayed 0.18 and 0.29 mm improvements of DIR to RO1 or RO2 compared to rigid registrations. The DIR to RO1 or RO2 MSD values were 1.44 and 1.67 mm greater than the RO1 to RO2 MSD, respectively, with means of 2.65 mm ($\pm$1.09 mm) and 2.88 mm ($\pm$1.35 mm).

As for the HD metric, this region showed the highest mean HD values compared to the other 2 regions. This region showed a better mean DIR to RO1 or RO2 (of all 3 metrics) than the RO1 to RO2. Specifically, the mean HD value was 15.73 mm ($\pm$7.60 mm) for DIR to RO1, compared to 18.92 mm ($\pm$20.47 mm) for RO1 to RO2; however, the mean rigid to RO1 was 13.72 mm ($\pm$5.66 mm). For further detail in regard to the 3 metrics for the region, refer to Figure 6.

Dice Similarity Coefficient Versus MSD Versus HD

When looking at all included contours, the rate at which DIR improved conformality to the physician contours from rigid registrations ranged from 50% to 65.6% across the 3 metrics. There was no specific trend shared between all 3 metrics in regard to which organs displayed more accurate contour registrations. Both DSC and MSD values shared the feature that the worst registrations were seen with seminal vesicles contours, while some of the most accurate registrations were seen with mandible contours. As for the HD metric, the best values were

seen with the spinal cord contours (in the H&N region) and the worst values were seen with the liver contours.

By way of the DSC, DIR improved the conformality from rigid registrations by a mean of 0.03 ($P < .05$) but still fell short of RO1 to RO2 conformality by 0.12 ($P < .005$). Based on MSD, the use of DIR showed a mean improvement of 0.33 mm ($P < .05$) but continued to fall short of RO1 to RO2 conformality by 1.23 mm ($P < .005$). With HD, the use of DIR lead to an increase by 0.25 mm, and it was greater than RO1 to RO2 by 1.99 mm.

Discussion

The use of DIR improved the accuracy of organ contours from rigid registrations alone a majority of the time when compared to manually drawn contours. Depending on the metric chosen, comparing the 3 regions' performances varied. The contours from patients with pancreatic cancer had the greatest agreement of DIR contours when using DSC. On the other hand, the use of MSD and HD showed the H&N region having the greatest DIR contour agreement.

As for the level of improvement via the use of DIR measured by each metric, DSC and MSD followed similar trends, while the HDs encountered challenges with several contours with high variability of ill-defined boundaries. The profound variation in trends seen among the 3 metrics continues to show the importance of inclusion of multiple metrics in order to evaluate conformality of DIR contours, as highlighted in
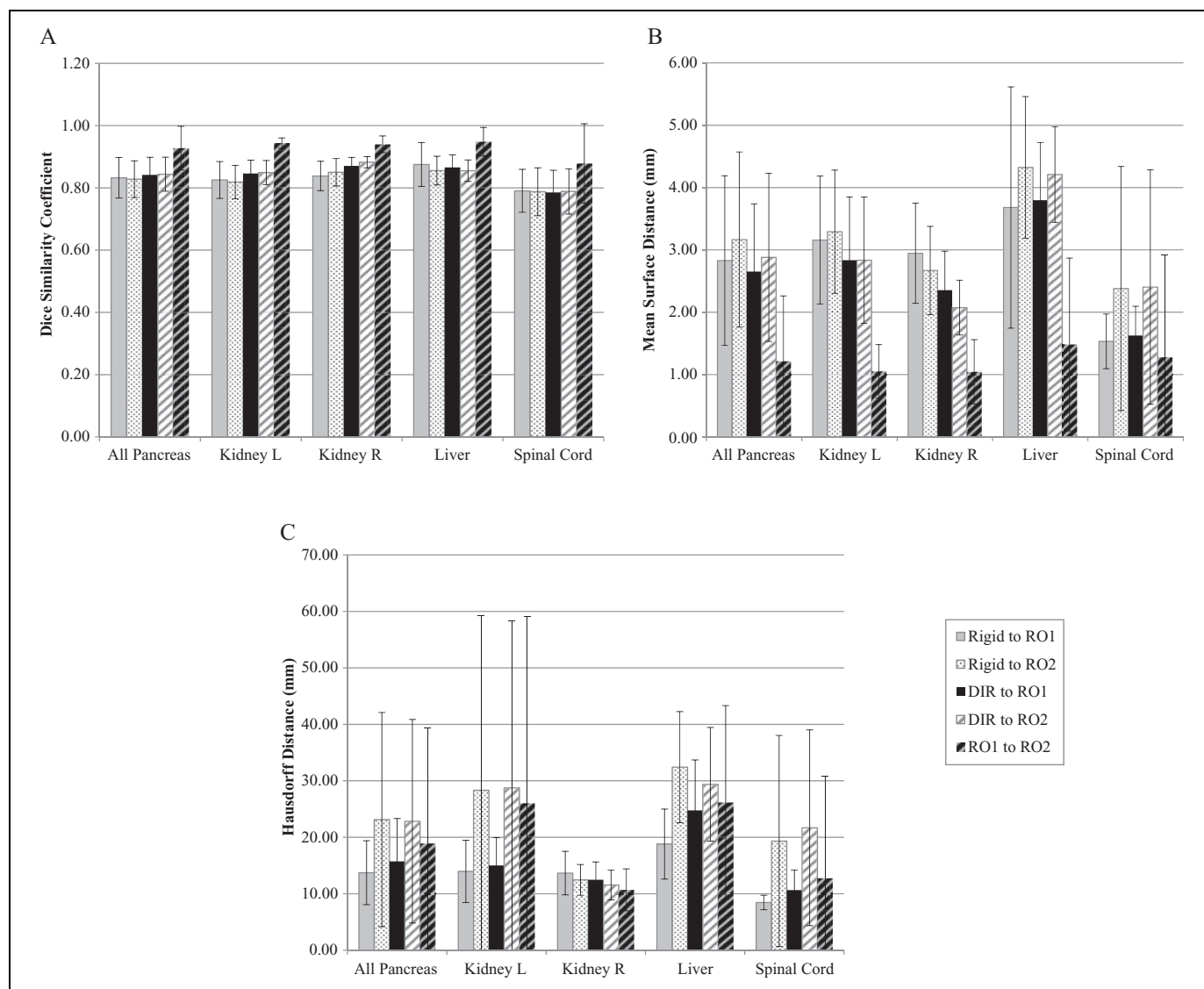
Figure 6. The DIR performance in the context of pancreas cancer. A, Mean DSC values for the pancreas region across the contoured organs. B, Mean MSD values of the pancreas region. C, Mean HD values of the pancreas region. DIR, indicates deformable image registration; DSC, Dice similarity coefficient; HD, Hausdorff distance; MSD, mean surface distance; RO1, radiation oncologist 1; RO2, radiation oncologist 2.

previous studies.[21] Larger structures with more defined boundaries, such as the mandible, tend to have more favorable DSC and MSD values. The higher Hounsfield units of bony structures may also contribute to a more favorable registration. The HD values, owing to the inherent nature of the metric, tend to be skewed by "outliers" of contour boundaries that are more common in larger and less defined organ contours seen in the abdomen. For instance, the liver had the highest HD values among all organs. The array of variations noticed among the metrics calls for guidelines for particular situations requiring each metric's clinical applicability.

In comparison with recent studies, Hvid et al showed DSC values >0.80 for all the OARs of the H&N region when performing CT to CBCT registrations.[22] This is a small improvement from the DSC values seen in our study, with a range of 0.70 to 0.85 for the OARs' DIR-to-physician comparisons.

A recent study by Ramadaan et al validating Varian's Eclipse DIR software for CT to CT registrations showed overall DSC of 0.84 for the H&N region and an improvement of 0.94 when reviewed by the treating RO.[23] Overall mean DSC of the H&N region for our study was 0.77; however, this was with the use of CBCT images, patients showing considerable anatomic change, and different organs of interest. Kumarasiri et al attained DSC values >0.85 for larger organs with clear boundaries in the H&N region, but the smaller organs with poorly defined boundaries had DSC values of 0.5 to 0.6.[18] This trend was seen in our study but to a much lesser extent. Larger, well-defined organs across the 3 regions had DSC values of 0.79 to 0.88, while the only organ that had DSC values less than 0.70 was the seminal vesicles (0.48-0.49). Overall, it appears that our results fall in line with some of the recent published studies.

There are some limitations to this study that need to be addressed. First, the study was limited to a sample size of 16 patients. The inclusion of 3 organ regions increases the possible applicability of the study; however, the different organs contoured between regions could hamper adequate comparisons. Second, the decision to have the physicians modify contours from a previous rigid registration could downplay the possible agreement of DIR contours to physician contours, as well as increase the physician-to-physician agreeability. Furthermore, the process of manually adjusting these rigid registered contours by the physicians was not timed. As for the DIR software itself, only 1 program was used during the study due to availability.

Another limitation related to the DIR software is the method registrations were compiled for image sets. There are several possible ways to attain more agreeable registrations, such as using single versus multiple deformations or modifying the region of interest prior to performing the registration. For this study, all registrations were assessed and approved by a medical physicist prior to applying the deformations to the contours. Future endeavors will assess the variations to these many DIR modalities.

A goal of this study was to quantify the improvement of DIR in the context of challenging patient anatomy. All of the patients selected for the study showed considerable changes over the course of treatment. Challenges that were identified, and sought after, included substantial tumor shrinkage in the H&N region, bladder or rectal filling disparities in the prostate region, and bowel and stomach filling in the pancreas region. Additionally, the differing modalities and qualities inherent to CBCT and CT images pose a challenge to CT–CBCT registrations. Given these less-than-ideal situations, DIR still showed improvement from the rigid registrations in a majority of the contours. Although the agreeability of DIR contours to manually drawn contours still appears to be inadequate, with its improved accuracy over rigid registrations and as previous studies have mentioned, the use of DIR in conjunction with manual adjustments could be a critical efficiency-improving method.[17]

Declaration of Conflicting Interests

Funding

References

1. Michalski JM, Gay H, Jackson A, Tucker SL, Deasy JO. Radiation dose–volume effects in radiation-induced rectal injury. *Int J Radiat Oncol*. 2010;76(3):S123-S129. doi:10.1016/j.ijrobp.2009.03.078.

2. Lotz HT, Remeijer P, van Herk M, et al. A model to predict bladder shapes from changes in bladder and rectal filling. *Med Phys*. 2004;31(6):1415-1423.

3. Xie Y, Chao M, Lee P, Xing L. Feature-based rectal contour propagation from planning CT to cone beam CT. *Med Phys*. 2008;35(10):4450-4459.

4. Landry JC, Yang GY, Ting JY, et al. Treatment of pancreatic cancer tumors with intensity-modulated radiation therapy (IMRT) using the volume at risk approach (VARA): employing dose-volume histogram (DVH) and normal tissue complication probability (NTCP) to evaluate small bowel toxicity. *Med Dosim*. 2002;27(2):121-129.

5. Harari PM, Song S, Tomé WA. Emphasizing conformal avoidance versus target definition for IMRT planning in head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2010;77(3):950-958. doi:10.1016/j.ijrobp.2009.09.062.

6. McBain CA, Henry AM, Sykes J, et al. X-ray volumetric imaging in image-guided radiotherapy: the new standard in on-treatment imaging. *Int J Radiat Oncol Biol Phys*. 2006;64(2):625-634. doi:10.1016/j.ijrobp.2005.09.018.

7. Chen T, Kim S, Goyal S, et al. Object-constrained meshless deformable algorithm for high speed 3D nonrigid registration between CT and CBCT. *Med Phys*. 2010;37(1):197-210.

8. Hardcastle N, van Elmpt W, De Ruysscher D, Bzdusek K, Tomé WA. Accuracy of deformable image registration for contour propagation in adaptive lung radiotherapy. *Radiat Oncol*. 2013;8:243. doi:10.1186/1748-717X-8-243.

9. Foroudi F, Wong J, Kron T, et al. Online adaptive radiotherapy for muscle-invasive bladder cancer: results of a pilot study. *Int J Radiat Oncol Biol Phys*. 2011;81(3):765-771. doi:10.1016/j.ijrobp.2010.06.061.

10. Nijkamp J, Pos FJ, Nuver TT, et al. Adaptive radiotherapy for prostate cancer using kilovoltage cone-beam computed tomography: first clinical results. *Int J Radiat Oncol Biol Phys*. 2008;70(1):75-82. doi:10.1016/j.ijrobp.2007.05.046.

11. Sarrut D. Deformable registration for image-guided radiation therapy. *Z Für Med Phys*. 2006;16(4):285-297.

12. Moulton CR, House MJ, Lye V, et al. Registering prostate external beam radiotherapy with a boost from high-dose-rate brachytherapy: a comparative evaluation of deformable registration algorithms. *Radiat Oncol*. 2015;10(1). doi:10.1186/s13014-015-0563-9.

13. Thor M, Petersen JBB, Bentzen L, Høyer M, Muren LP. Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer. *Acta Oncol*. 2011;50(6):918-925. doi:10.3109/0284186X.2011.577806.

14. Hou J, Guerrero M, Chen W, D'Souza WD. Deformable planning CT to cone-beam CT image registration in head-and-neck cancer. *Med Phys*. 2011;38(4):2088-2094. doi:10.1118/1.3554647.

15. Veiga C, Lourenço AM, Mouinuddin S, et al. Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm. *Med Phys*. 2015;42(2):760-769. doi:10.1118/1.4905050.

16. Pukala J, Meeks SL, Staton RJ, Bova FJ, Mañon RR, Langen KM. A virtual phantom library for the quantification of deformable image registration uncertainties in patients with cancers of the head and neck. *Med Phys*. 2013;40(11):111703. doi:10.1118/1.4823467.

17. Perna L, Sini C, Cozzarini C, et al. Deformable registration-based segmentation of the bowel on megavoltage CT during pelvic

radiotherapy. *Phys Med*. 2016;32(7):898-904. doi:10.1016/j.
ejmp.2016.06.009.

18. Kumarasiri A, Siddiqui F, Liu C, et al. Deformable image regis-
tration based automatic CT-to-CT contour propagation for head
and neck adaptive radiotherapy in the routine clinical setting. *Med
Phys*. 2014;41(12):121712. doi:10.1118/1.4901409.

19. Dice LR. Measures of the amount of ecologic association between
species. *Ecology*. 1945;26(3):297-302. doi:10.2307/1932409.

20. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing
images using the Hausdorff distance. *IEEE Trans Pattern Anal
Mach Intell*. 1993;15(9):850-863. doi:10.1109/34.232073.

21. Varadhan R, Karangelis G, Krishnan K, Hui S. A framework for
deformable image registration validation in radiotherapy clinical
applications. *J Appl Clin Med Phys*. 2013;14(1):4066.

22. Hvid CA, Elstrøm UV, Jensen K, Alber M, Grau C. Accuracy of
software-assisted contour propagation from planning CT to cone
beam CT in head and neck radiotherapy. *Acta Oncol*. 2016;
55(11):1324-1330. doi:10.1080/0284186X.2016.1185149.

23. Ramadaan IS, Peick K, Hamilton DA, et al. Validation of Var-
ian's SmartAdapt® deformable image registration algorithm for
clinical application. *Radiat Oncol*. 2015;10(1):73. doi:10.1186/
s13014-015-0372-1.