Quality Assurance Assessment of Diffusion-Weighted and T2-Weighted Magnetic Resonance Imaging Registration and Contour Propagation for Head and Neck Cancer Radiotherapy

Mohamed A. Naser[a]*, Kareem A. Wahid[a]*, Sara Ahmed[a], Vivian Salama[a], Cem Dede[a], Ruitao Lin[b], Brigid McDonald[a], Travis C. Salzillo[a], Renjie He[a], Yao Ding[c], Moamen Abobakr Abdelaal[a], Daniel Thill[d], Nicolette O'Connell[d], Virgil Willcut[d], John P. Christodouleas[d], Stephen Y Lai[e], Clifton D. Fuller[a]**, Abdallah S. R. Mohamed[a]**

[a]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

[b]Department, of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

[c]Department, of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

[d]Elekta AB, Stockholm, Sweden.

[e]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

*co-first authors.

**co-corresponding authors.

Corresponding authors: Clifton D. Fuller. Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA. Email: cdfuller@mdanderson.org. Phone number: 713-745-4404. Abdallah S.R. Mohamed. Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA. Email: asmohamed@mdanderson.org. Phone number: 713-745-4092. Postal Address: The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, 77030, USA.

Abstract:

Background/Purpose: Adequate image registration of anatomic and functional MRI scans is necessary for MR-guided head and neck cancer (HNC) adaptive radiotherapy planning. Despite the quantitative capabilities of diffusion-weighted imaging (DWI) MRI for treatment plan adaptation, geometric distortion remains a considerable limitation. Therefore, we systematically investigated various deformable image registration (DIR) algorithms to co-register DWI and T2-weighted (T2W) images.

Materials/Methods: We compared post-acquisition registration algorithms from three software packages (ADMIRE, Velocity, and 3D Slicer) applied to T2W and DWI MRI images in twenty HNC patients. In addition, we investigated implicit rigid registration (no algorithm applied) as a control comparator. Ground truth segmentations of radiotherapy structures (tumor and organs at risk) were generated by a physician expert on both image sequences. Three additional experts provided segmentations for five cases for interobserver variability studies. For each registration approach, structures were propagated from T2W to DWI images. These propagated structures were then compared with ground truth DWI structures using the Dice similarity coefficient (DSC), false-negative DSC, false-positive DSC, surface DSC, 95% Hausdorff distance, and mean surface distance.

Results: 19 left submandibular glands, 18 right submandibular glands, 20 left parotid glands, 20 right parotid glands, 20 spinal cords, 9 brainstems, and 12 tumors were delineated. ADMIRE, the atlas-based auto segmentation DIR algorithm, demonstrated improved performance over implicit rigid registrations for most comparison metrics and structures (Bonferroni-corrected $p < 0.05$), while Velocity and 3D Slicer algorithms did not. Moreover, the ADMIRE methods significantly

improved performance in individual and pooled analysis compared to all other methods. Interobserver variability analysis revealed no significant difference between observers ($p > 0.05$).

Conclusions: Certain deformable registration software packages, such as those provided by ADMIRE, may be favorable for registering T2W and DWI images. These results are important to ensure the appropriate selection of registration strategies for MR-guided radiotherapy.

Keywords: deformable image registration, magnetic resonance, adaptive radiotherapy, quality assurance.

Abbreviations: T2-weighted (T2W), diffusion-weighted imaging (DWI), Radiation therapy (RT), head and neck cancer (HNC), organs at risk (OAR), deformable image registration (DIR).

1. Introduction:

Radiation therapy (RT) is an essential treatment modality for head and neck cancer (HNC) [1]. Conventionally, RT has relied on radiographic images to enable pre-treatment segmentation of target volumes and nearby organs at risk (OAR) to plan intensity-modulated doses [2,3]. However, throughout RT, the dynamic changes in target volumes and OARs and patient-specific changes (e.g., weight loss) can lead to unintended doses of radiation to OARs and subsequent debilitating side effects [4]. These potential unintended doses are particularly relevant for HNC because the head and neck region is home to various complex, highly radiosensitive structures and tissue interfaces that can drastically change during RT [4,5].

Image-guided RT, during which radiation dose can be administered in tandem with onboard imaging, has become a promising alternative to intensity-modulated RT, in part due to increasingly ubiquitous image-guided technology, such as MR-Linac devices [6,7]. MR-guided

treatment also affords the ability to capture distinct patient anatomy with varying contrasts via weighted sequence acquisitions, such as T2-weighted (T2W) images, and functional information, such as through diffusion-weighted imaging (DWI). DWI has shown particular benefit in aiding treatment adaptation through improved detection of target volumes and assessment of treatment response [8]. Therefore, combined T2W and DWI acquisition enable the gathering of anatomic and functional information that can be used for adaptive MR-guided personalized RT.

Anatomical and functional sequences acquired in the same imaging session for MR-guided treatment often have minimal variation in patient position and geometry between sequence acquisitions. However, these multisequence acquisitions can be misaligned by motion artifacts from respiration or swallowing [4], susceptibility artifacts, chemical shift artifacts, ghosting artifacts [8], and geometric distortions [9]. Post-acquisition image registration, the process by which homologous image voxels from multi-temporal or multi-modal image sets are mapped to each other [10,11], is an important approach to align anatomical and functional sequences. Rigid image registration involves global matching between image sets, while deformable image registration (DIR) uses optimization algorithms to adjust image transformation models. Most implementations of DIR involve a transformation that establishes a geometric correspondence between fixed and moving images, an objective function, and an optimization approach to maximize the similarity between images [12–14]. Importantly, even minor differences in patient anatomy can result in devastating dose administration in HNC [4,15], highlighting the need for consistent image co-registration when propagating segmentations of target volumes and OARs for radiotherapy treatment planning. Therefore, determining the impact of post-acquisition registration techniques (i.e., DIR) on multisequence MRI acquisitions is crucial for MR-guided treatment of HNC.

While we have previously investigated intra-modality CT to CT registration [16] and inter-modality CT to MRI registration [17], to our knowledge, there are no studies that investigate registration techniques for intra-acquisition MRI in HNC. Therefore, to facilitate further development and optimization of MR-guided RT adaptive planning technologies, we systematically analyzed registration methods in T2W and DWI MRI sequences acquired during the same imaging session.

2. Methods:

We developed a quality assurance workflow for evaluating and benchmarking the performance of different image registration algorithms for T2W and DWI images (described in subsection 2.2) of 20 HNC patients (2.1). We evaluated different DIR algorithms provided by ADMIRE and Velocity as well as B-spline and affine algorithms by 3D Slicer (2.3). The deformation vector field generated by each algorithm was used to propagate the manually segmented structures (2.2) from T2W to DWI images. The structures propagated by different algorithms to DWI images were compared to the ground truth segmentations for performance evaluation (2.4).

2.1. Patient Characteristics: Twenty patients with HNC who had undergone RT in a clinical trial (NCT03145077) were included in this analysis. All clinical and imaging data were generated between May 30, 2017 and April 1, 2019 and were retrospectively collected under a HIPAA-compliant protocol (PA16-0302) that was approved by The University of Texas MD Anderson Cancer Center's institutional review board. All patients provided study-specific informed consent. The median patient age was 54 years, with a male predominance (80%). Primary tumor sites included the oropharynx, nasopharynx, and oral cavity. Full patient clinical and demographic characteristics are summarized in **Table 1**.

Table 1. Patient clinical and demographic characteristics.

Characteristic	Value
Age (median, range)	54 (32-77)
Sex	
Male	16
Female	4
Race	
Asian	1
White/Caucasian	19
Tumor subsite	
Base of tongue	10
Tonsil	5
Buccal mucosa	1
Floor of mouth	1
Nasal cavity	1
Nasopharynx	1
Unknown	1
T-category	
T0	1
T1	5
T2	7
T3	2
T4	5
N-category	
Nx	1
N1	11
N2a	1
N2b	6
N2c	1
AJCC stage	
I	2
II	6
III	2
IVA	9
IVB	1

Unless otherwise indicated, data shown correspond to patient number counts. AJCC, American Joint Committee on Cancer.

2.2. Imaging Data: Pre-RT T2W and DWI MRI sequences in Digital Imaging and Communications in Medicine (DICOM) format for each of the 20 patients were curated from our imaging databases. T2W images and DWI images with a b value of 0 were collected in the same imaging session while the patient was immobilized in a thermoplastic mask using a 1.5

Tesla Siemens MRI simulator. Characteristics of the imaging sequences are shown in **Table 2**. For each image set (T2W image and DWI image), ground truth segmentations for the left and right submandibular glands, left and right parotid glands, cervical spinal cord, brainstem, and primary gross tumor volume were manually generated by a trained physician expert (radiologist with > 5 years of experience in HNC). In addition, in a subset of five cases, segmentations for all structures in both sequences were manually generated by three additional separate observers (two physicians and one medical student) for interobserver variability analysis. All segmentations were generated in Velocity AI (v.3.0.1; Varian Medical Systems; Palo Alto, CA, USA) in DICOM RT structure format. The anonymized image sets and structure files are publicly available online through Figshare (10.6084/m9.figshare.17162435).

Table 2. MRI sequence acquisition parameters.

Acquisition Parameter	T2W	DWI
Repetition time (ms)	4800.00	5400.00
Echo time (ms)	80.00	50.00
Echo train length	15	15
Flip angle (°)	180	120
Slice thickness (mm)	2.00	4.00
In-plane resolution (mm)	0.50	2.00
Acquisition matrix	256x230	128x128
Pixel bandwidth (Hz/px)	300.00	1220.00
Number of averages	1	8

T2W, T2-weighted magnetic resonance imaging; DWI, diffusion-weighted magnetic resonance imaging.

2.3. Image Registration: For this analysis, we investigated several image registration algorithms from different imaging software packages: ADMIRE (v.3.29; Elekta AB; Stockholm, Sweden), Velocity AI (v.3.0.1; Varian Medical Systems; Palo Alto, CA, USA), and 3D Slicer (v.4.11; open-

source [18]). A total of six image registration methods were investigated in this study: ADMIRE, ADMIRE_Skiplinear, Velocity, Velocity_Skiplinear, B-spline, and Affine. From ADMIRE and Velocity AI, we used a rigid image registration followed by a DIR (ADMIRE, Velocity) and a DIR alone (ADMIRE_Skiplinear, Velocity_Skiplinear). The ADMIRE DIR algorithms utilize an atlas-based approach with head pose correction, dense mutual-information, and final refinement using a deformable surface model [19]. From 3D Slicer, we used a B-spline registration (B-spline) — a nonlinear parametric model commonly used for DIR [20] — and an Affine registration (Affine). For all cases, the DWI image was used as the fixed image, and the T2W image was used as the moving image. As a control comparator for all cases, we also analyzed the raw images with no post-acquisition registration applied, i.e., an implicit rigid registration (Rigid). After the registration process, the ground truth segmentations from the T2W images were propagated to the corresponding DWI images to generate propagated structures (**Figure 1A-H**). These propagated structures were then compared to the ground truth structures on the DWI image in the subsequent analysis. Before the analysis, all images and structure files were transformed into Neuroimaging Informatics Technology Initiative format in 3D Slicer. Moreover, to maintain adequate comparisons between structures generated on T2W images and DWI images, all images and structures were cropped to the image with the smaller field of view, e.g., DWI image. Finally, because there were small variations in the inferior and superior slices of the cervical spinal cord and brainstem structures between image sets, we cropped these structures so that the heights of the propagated segmentation and ground truth segmentation were equal (**Figure 1I)**. We utilized a similar process for the interobserver variability analysis, in which segmentations were propagated from T2W to DWI images and then compared to ground truth segmentations generated by each individual observer.

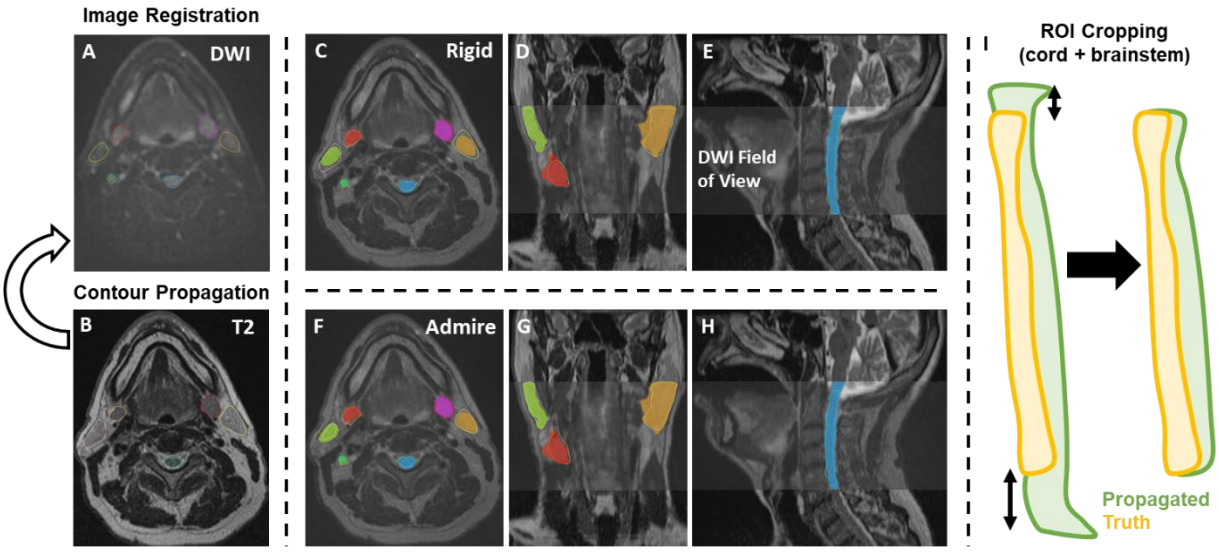

Figure 1. Study workflow. Contours are propagated from the moving image (**B**, T2-weighted image [T2]) to the fixed image (**A**, diffusion-weighted image [DWI]) for each registration algorithm. **C-E** and **F-H** show propagated and ground truth structures for the implicit rigid and ADMIRE_Skiplinear approaches, respectively. For the spinal cord and brainstem, the regions of interest (ROI) were also cropped so that the height of the propagated segmentation and ground truth segmentation were equal (**I**).

2.4. Statistical Analysis: Several evaluation metrics were used to compare the propagated structure sets after registration to the ground truth structures delineated on the DWI images. For each individual structure, the following evaluation metrics were calculated between the propagated structure sets after registration and the ground truth structures delineated on the DWI images: Dice similarity coefficient (DSC), false-positive DSC (FP-DSC), false-negative DSC (FN-DSC), surface DSC (S-DSC), 95% Hausdorff distance (95% HD), and mean surface distance (MSD). For S-DSC, a tolerance of 2.5 mm was selected as a suitable tolerance from previous studies [21,22]. Surface distance metrics were calculated using the surface-distance Python package [21], and all other metrics were calculated in ADMIRE (v.3.29; Elekta AB;

Stockholm, Sweden). After performing a Shapiro-Wilk test [23], we found that our data were not normally distributed ($p<0.05$). Therefore, we used nonparametric statistical tests for our analysis. For each metric and each structure, we compared registration algorithms against the implicit rigid registration using one-sided Wilcoxon signed-rank tests (alternative hypothesis of greater than the null hypothesis for DSC and S-DSC and alternative hypothesis of less than the null hypothesis for FP-DSC, FN-DSC, 95% HD, and MSD) with Bonferroni adjustments for multiple comparisons [24]. Similarly, we pooled metrics for OARs for sub-analysis and performed pair-wise analysis using previously described Wilcoxon signed-rank tests with Bonferroni corrections. For interobserver variability analysis of registration algorithms, we implemented a Kruskal-Wallis one-way analysis of variance test [25] for all four observers across all structures and evaluation metrics. For all statistical analyses, p-values less than 0.05 were considered significant. All statistical analyses were performed in Python v.3.7 [26].

3. Results:

3.1. Comparison of Registration Methods: 19 left submandibular glands, 18 right submandibular glands, 20 left parotid glands, 20 right parotid glands, 20 cervical spinal cords, 9 brainstems, and 12 primary tumors in both T2W and DWI images were used in the analysis. For each algorithm, most structures had similar performance across the multiple metrics except for the cervical spinal cord (**Figure 2**); we therefore performed additional analysis to investigate spinal cord structures in individual cases in **Appendix A**. Compared to the structures generated by the implicit rigid registration, almost all structures demonstrated an improvement with the ADMIRE based methods, and almost all structures worsened with the Velocity based methods; most 3D Slicer based methods had similar results those of the implicit rigid registration (**Figure 2**). Specifically, the ADMIRE_Skiplinear method offered the best overall performance, with significant improvements ($p<0.05$ on one sided Wilcoxon signed rank test) in 5/6 metrics for the

left submandibular gland, 2/6 metrics for the right submandibular gland, 5/6 metrics for the left parotid gland, 4/6 metrics for the right parotid gland, 4/6 metrics for the spinal cord, 1/6 metrics for the brainstem, and 3/6 metrics for the tumor (**Figure 3**). When metrics were pooled across structures, similar trends emerged where the ADMIRE and ADMIRE_Skiplinear methods demonstrated the best performance compared to the other methods, with DSC gains over the implicit rigid registration of up to 0.04 in the OARs and tumor (**Table 3**). Moreover, pair-wise comparisons of pooled OAR structures and the tumor demonstrated that the ADMIRE based methods significantly improved performance ($p<0.05$) over other methods, and most methods outperformed the Velocity based methods (**Figure 4**).

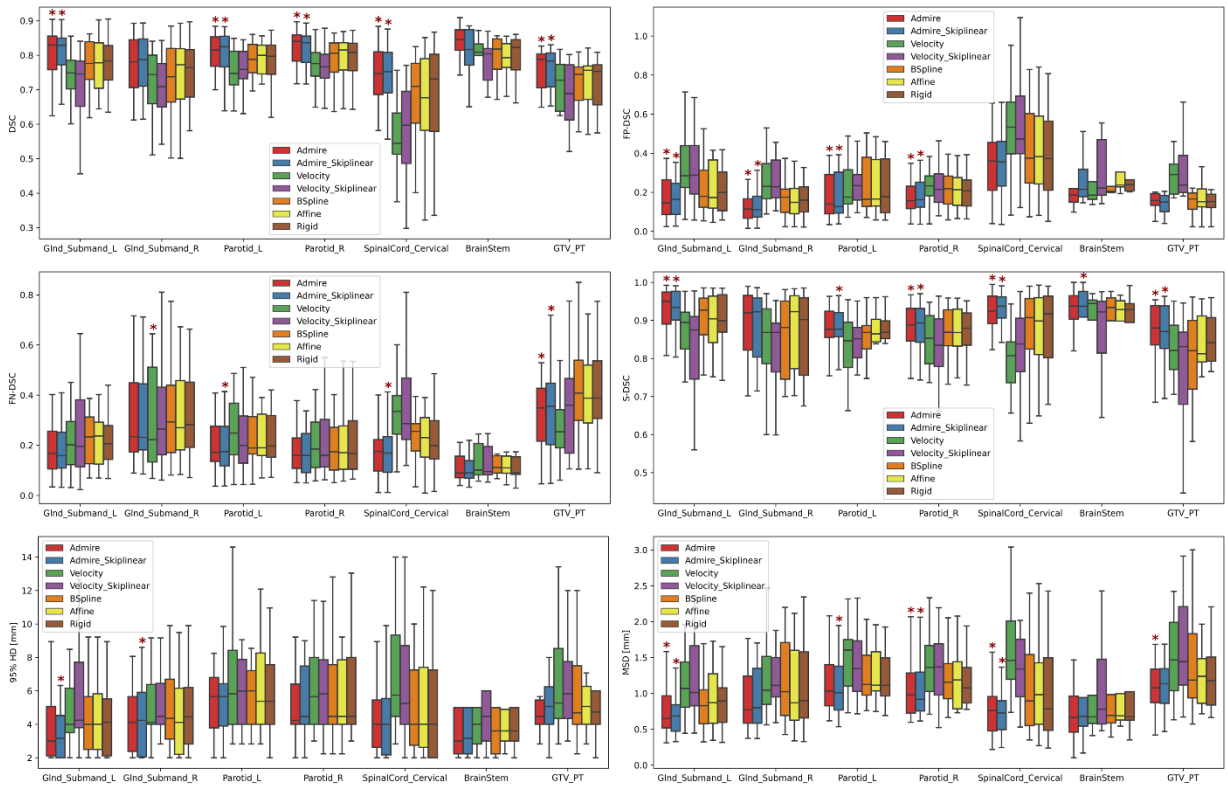

Figure 2. Box plots of evaluation metrics for each structure according to the registration method. Asterisks indicate a significant improvement between the registration method and

implicit rigid registration. Glnd_Submand, submandibular gland; L, left; R, right; GTV_PT, primary gross tumor volume; DSC, Dice similarity coefficient; FP-DSC, false-positive Dice similarity coefficient; FN-DSC, false-negative Dice similarity coefficient; S-DSC, surface Dice similarity coefficient; 95% HD, 95% Hausdorff distance; MSD, mean surface distance.

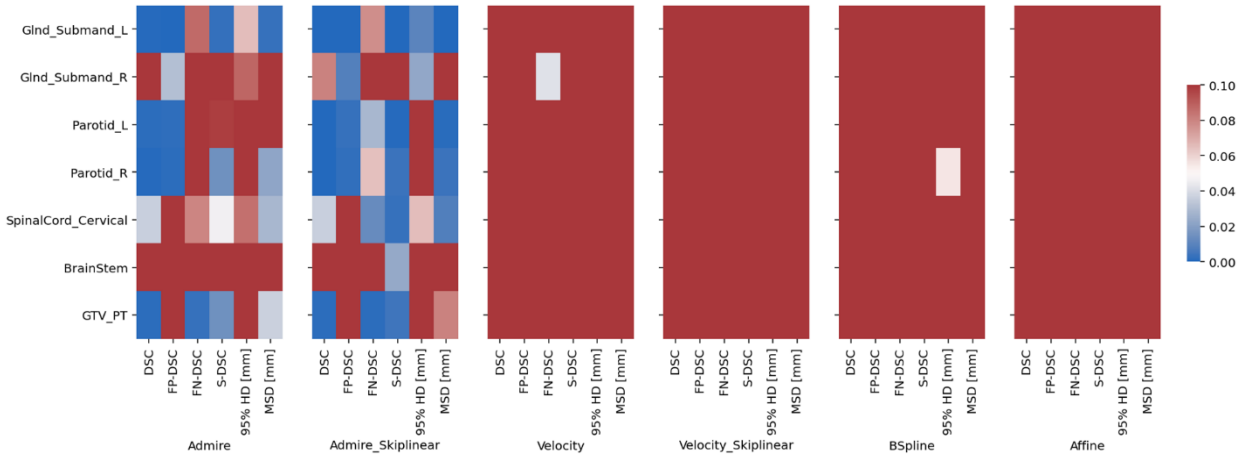

Figure 3. Heatmap of Bonferroni corrected p-values for one-way Wilcoxon-signed rank tests between various registration methods and implicit rigid registration indicating significant improvement across evaluation metrics and structures. Blue colors correspond to significant p-values ($p<0.05$) while red colors correspond to non-significant values ($p>0.05$). Glnd_Submand, submandibular gland; L, left; R, right; GTV_PT, primary gross tumor volume; DSC, Dice similarity coefficient; FP-DSC, false-positive Dice similarity coefficient; FN-DSC, false-negative Dice similarity coefficient; S-DSC, surface Dice similarity coefficient; 95% HD, 95% Hausdorff distance; MSD, mean surface distance.

Table 3. Evaluation metrics (mean ± standard deviation) across pooled structures according to each registration method.

	Metric	ADMIRE	ADMIRE_ Skiplinear	Velocity	Velocity_ Skiplinear	Affine	B-Spline	Rigid
Normal Tissue (OAR)	DSC	0.79±0.09	0.79±0.08	0.71±0.13	0.7±0.15	0.75±0.11	0.75±0.11	0.75±0.11
	FP-DSC	0.22±0.17	0.22±0.16	0.32±0.2	0.33±0.22	0.26±0.18	0.27±0.18	0.26±0.18
	FN-DSC	0.21±0.14	0.2±0.14	0.26±0.17	0.28±0.2	0.24±0.15	0.24±0.16	0.24±0.15
	S-DSC	0.9±0.07	0.9±0.07	0.84±0.1	0.83±0.13	0.88±0.09	0.87±0.09	0.88±0.09
	95% HD (mm)	5.16±3.83	5.19±3.74	6.49±4.21	6.64±4.27	5.76±4.03	5.64±3.75	5.71±3.98
	MSD (mm)	0.95±0.51	0.95±0.49	1.36±0.62	1.45±0.94	1.14±0.58	1.14±0.58	1.13±0.59
Tumor	DSC	0.75±0.08	0.75±0.08	0.68±0.14	0.62±0.21	0.71±0.1	0.7±0.11	0.71±0.1
	FP-DSC	0.17±0.07	0.16±0.07	0.31±0.14	0.36±0.24	0.17±0.09	0.17±0.09	0.17±0.08
	FN-DSC	0.34±0.18	0.35±0.17	0.33±0.25	0.4±0.27	0.41±0.17	0.43±0.2	0.41±0.18
	S-DSC	0.87±0.08	0.87±0.08	0.82±0.1	0.75±0.18	0.83±0.1	0.81±0.11	0.84±0.1
	95% HD [mm]	6.04±4.03	6.23±3.89	6.81±3.85	6.66±3.5	6.67±4.28	7.37±5.44	6.83±5.19
	MSD [mm]	1.16±0.52	1.16±0.48	1.49±0.57	1.87±1.28	1.35±0.62	1.45±0.73	1.33±0.65

OAR, organs at risk; DSC, Dice similarity coefficient; FP-DSC, false-positive Dice similarity coefficient; FN-DSC, false-negative Dice similarity coefficient; S-DSC, surface Dice similarity coefficient; 95% HD, 95% Hausdorff distance; MSD, mean surface distance.

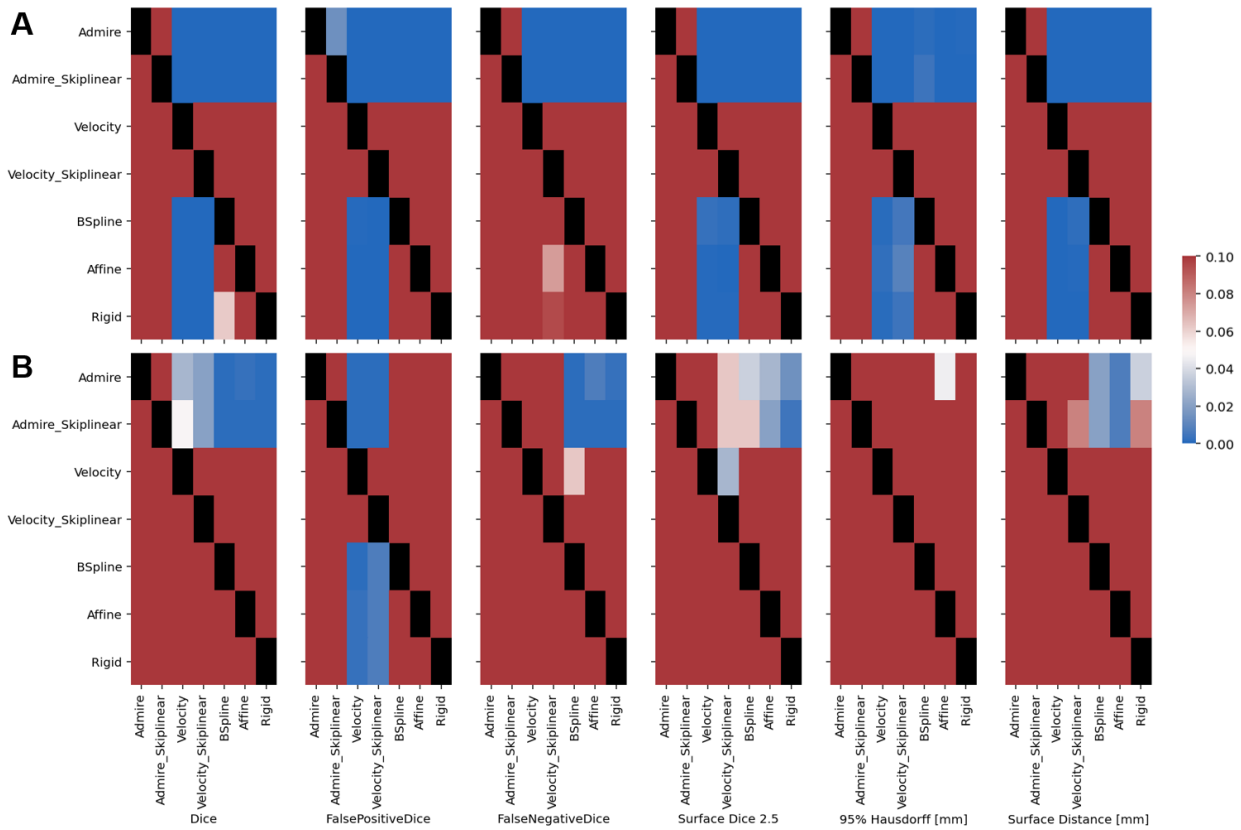

Figure 4. Heatmap of Bonferroni corrected p-values for one-way Wilcoxon-signed rank tests for pair-wise comparisons between various registration methods for (**A**) all pooled organ at risk structures and (**B**) tumor indicating significant improvement of method in row vs. method in column. Blue colors correspond to significant p-values ($p<0.05$) while red colors correspond to non-significant values ($p>0.05$). Comparisons between the same algorithm (diagonal entries) are blacked out.

3.2. Interobserver Variability Analysis: We performed an interobserver variability analysis to determine if there were any significant differences between observers for a given registration method. Metric value comparisons between all observers were non-significant for all structures (**Figure 5**); therefore, our study had no major interobserver variability.

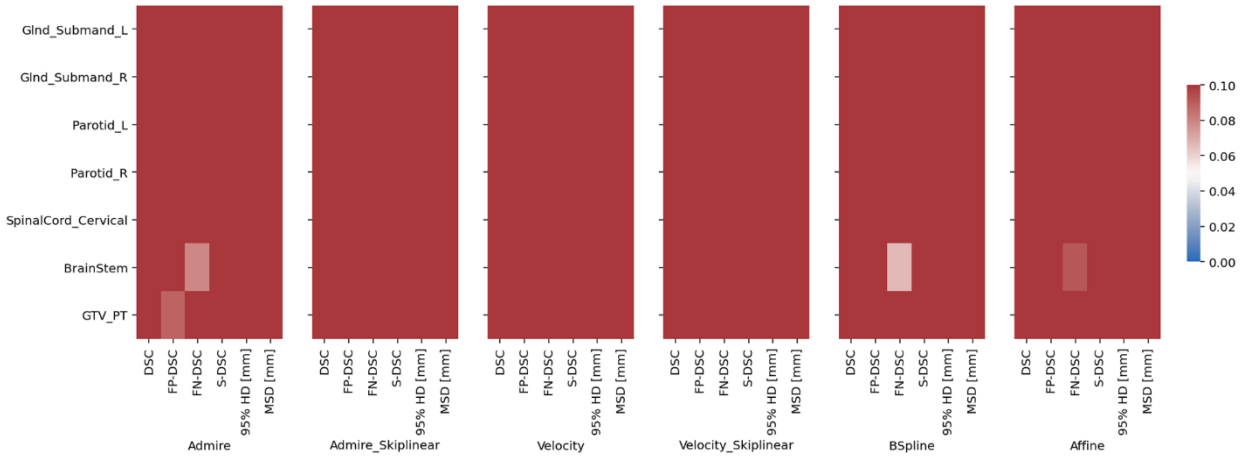

Figure 5. Heatmap of p-values for Kruskal-Wallis one-way analysis of variance test for all four observers across all structures and evaluation metrics. Blue colors correspond to significant p-values ($p<0.05$) while red colors correspond to non-significant values ($p>0.05$). Glnd_Submand, submandibular gland; L, left; R, right; GTV_PT, primary gross tumor volume; DSC, dice similarity coefficient; FP-DSC, false-positive dice similarity coefficient; FN-DSC, false-negative dice similarity coefficient; S-DSC, surface dice similarity coefficient; 95% HD, 95% Hausdorff distance; MSD, mean surface distance.

4. Discussion

In this study, we systematically analyzed a variety of registration algorithms and compared them to implicitly registered images from multisequence MRI acquisitions for image-guided treatment applications. Our results highlight that specific registration algorithms can improve upon implicitly registered image registration quality, as shown by measuring the similarity of propagated ground truth segmentations from T2W images to DWI images compared to ground truth segmentations on DWI images (**Figures 2 and 3**).

The best overall results were obtained using the ADMIRE software's deformable registration algorithm ADMIRE_Skiplinear (**Table 3**), with most metrics and structure combinations having better performance than the implicit rigid registration (**Figure 3**). Interestingly, the application of an additional rigid registration before the ADMIRE algorithm (ADMIRE) often decreased performance (**Figure 2**), though not significantly (**Figure 4**). This minorly decreased performance may be secondary to slight differences in alignment imposed by the added rigid registration step. While we tested other deformable methods (i.e., Velocity and 3D Slicer), with the exception of a few outliers, they did not demonstrate significantly improved performance for most metric and structure combinations when compared to the implicit rigid registration. Moreover, the Velocity based methods were often worse than the implicit rigid registration (**Figure 4**), which may be due to the DIR algorithmic implementation being unable to accommodate large variations in intensity domains of the T2W and DWI images. Importantly, almost all structures of interest individually and on pooled analysis showed increased DSC for the ADMIRE based methods (**Table 3**). This result indicates that the ADMIRE based methods provide significantly improved volumetric overlap, which may warrant their use during intra-acquisition MRI sequences for MR-guided treatment. It is worth noting the spinal cord is especially sensitive to distortion-causing artifacts [27], making it a particularly challenging structure to co-register adequately. While the general performance for the spinal cord was lower than that of other structures, the ADMIRE based methods were still able to offer significantly improved performance compared to the implicit rigid registration (**Figures 2 and 3)**; cases with lower performance tended to have a larger degree of spinal curvature than cases with higher performance (**Appendix A**). Therefore, while the ADMIRE based methods should still be preferred over implicit rigid registration, special caution should be used in quality assurance of these algorithms when used for spinal cord segmentations. Notably, all estimated metrics between any registration algorithm and the implicit rigid registration showed no significant

differences using segmentations generated by different observers (**Figure 5**), indicating that our data are not confounded by interobserver variability.

While several previous studies have investigated the relative performance of registration algorithms in various anatomical sites [28–30], there is a general lack of investigations of head and neck imaging. However, a few recent important studies have investigated registration quality assessment in head and neck imaging using radiotherapy structure analysis similar to our current study [16,17]. For example, Mohamed et al. [16] investigated the registration quality of diagnostic CT to simulation CT in HNC where images were acquired at different time points and with different scan settings and found that certain DIR methods demonstrated improved performance over a control group (rigid registration) for OAR and target conformance for most comparison metrics, similar to our study. Oppositely, Kiser et al. [17] showed that for CT and T2W MRI scans acquired with standard treatment immobilization techniques, MRI to CT DIR was not superior to rigid registration, with neither technique producing clinically satisfactory results (DSCs of 0.62 - 0.65). Importantly, the ADMIRE algorithms investigated in our study produce potentially clinically meaningful results as we observe significant performance gains across various structures that may impact MR-guided treatments.

To date, no anatomical to functional MRI registration studies have been performed for HNC. However, intra-acquisition MRI registration techniques have been investigated in other anatomical sites. Specifically, several studies have compared registration techniques for various MRI sequences in the prostate [31–33]. For example, Buerger et al. compared the performance of five state-of-the-art DIR image registration techniques for accurate image fusion of DWI with T2W images and found fast elastic image registration provided improved performance

compared to other deformable techniques such as B-spline and Demons [32]. This result was further echoed in Eriksson et al., which confirmed that fast elastic image registration was the best technique for T1-weighted to T2W anatomic sequence registration [33]. Our results are consistent with these observations that selecting appropriate deformable techniques offers significantly improved performance for intra-acquisition registration.

There are several limitations to our study. We limited our analysis of intra-acquisition registration techniques in MRI to T2W and DWI sequences since these are the most germane to current MR-guided RT applications. However, several additional sequences can be studied to investigate these phenomena. For example, a study by Chen et al. in brain imaging investigated the best DWI-derived scalar images for T1-weighted to DWI image co-registration [34]. The authors discovered that anisotropic power images provide the most consistent registration compared to all other images, with b0 images offering average performance. In this study, we only tested b0 images, which were readily available and common for DWI workflows in HNC. Future iterations of this study should investigate other DWI-derived scalar images. Additionally, we have investigated a few select volumetric overlap and surface distance metrics since these are the most ubiquitous metrics used in evaluating segmentation quality for RT applications [35]. However, dosimetric studies on OARs and target volumes may be warranted to determine the ultimate impact of registration quality on a RT plan. Finally, we have limited our analysis to intra-acquisition images collected during the same image acquisition session. However, for MR-guided RT applications, registration techniques are also relevant for images taken at different time points. Therefore, future studies should investigate these registration techniques applied to different imaging time points in an MR-guided RT workflow.

5. Conclusions

In summary, this is the first study to investigate intra-acquisition MRI registration quality in HNC patients. We identify a deformable registration technique from the ADMIRE software package that offers the most significant gains in registration quality for T2W to DWI image registration compared to other methods. Our results are a crucial first step towards registration quality assurance for MR-guided treatment approaches that implement multi-sequence acquisitions combining anatomical and functional imaging.

Acknowledgments

The authors thank Ann Sutton, Scientific Editor, and Ashli Nguyen-Villarreal, Associate Scientific Editor, in the Research Medical Library at The University of Texas MD Anderson Cancer Center, for editing this article. The authors also acknowledge the following people for their contributions to the NIH-funded academic-industrial partnership grant (R01DE028290) that funded this work and for their general support and feedback regarding this project: Spencer Marshall, Hafid Akhiat, Michel Moreau, Nathan Cho, Edyta Bubula-Rehm, Chunhua Men, and Etienne Lessard of Elekta and Alex Dresner of Philips.

Funding Statement and Declaration of Interests

This work was supported by the National Institutes of Health (NIH) through a Cancer Center Support Grant (P30-CA016672-44). K.A. Wahid and T.C. Salzillo are supported by training fellowships from The University of Texas Health Science Center at Houston Center for Clinical and Translational Sciences TL1 Program (TL1TR003169). K.A. Wahid is also supported by the American Legion Auxiliary Fellowship in Cancer Research and a NIDCR F31 fellowship (1 F31 DE031502-01). B.A. McDonald receives research support from an NIH NIDCR Award (F31DE029093) and the Dr. John J. Kopchick Fellowship through The University of Texas MD Anderson UTHealth Graduate School of Biomedical Sciences. C.D. Fuller received funding from

an NIH NIDCR Award (1R01 DE025248-01/R56 DE025248) and Academic-Industrial Partnership Award (R01 DE028290); the National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679); the NIH Big Data to Knowledge (BD2K) Program of the National Cancer Institute (NCI) Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01 CA214825); the NCI Early Phase Clinical Trials in Imaging and Image-Guided Interventions Program (1R01 CA218148); the NIH/NCI Cancer Center Support Grant (CCSG) Pilot Research Program Award from the UT MD Anderson CCSG Radiation Oncology and Cancer Imaging Program (P30 CA016672); the NIH/NCI Head and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50 CA097007); and the National Institute of Biomedical Imaging and Bioengineering (NIBIB) Research Education Program (R25 EB025787). He has received direct industry grant support, speaking honoraria, and travel funding from Elekta AB.

References

[1] Koyfman SA, Brizel DM, Posner MR. General principles of radiation therapy for head and neck cancer. UpToDate; Post, TW, Ed; UpToDate Waltham, MA, USA 2018. https://www.uptodate.com/contents/general-principles-of-radiation-therapy-for-head-and-neck-cancer (accessed August 20, 2021).

[2] Bortfeld T. IMRT: a review and preview. Phys Med Biol 2006;51:R363.

[3] Marta GN, Silva V, de Andrade Carvalho H, de Arruda FF, Hanna SA, Gadia R, et al. Intensity-modulated radiation therapy for head and neck cancer: systematic review and meta-analysis. Radiother Oncol 2014;110:9–15.

[4] Mali SB. Adaptive radiotherapy for head neck cancer. J Maxillofac Oral Surg 2016;15:549–54.

[5] Morgan HE, Sher DJ. Adaptive radiotherapy for head and neck cancer. Cancers Head Neck 2020;5:1. https://doi.org/10.1186/s41199-019-0046-z.

[6] Pollard JM, Wen Z, Sadagopan R, Wang J, Ibbott GS. The future of image-guided radiotherapy will be MR guided. Br J Radiol 2017;90:20160667.

[7] Choudhury A, Budgell G, MacKay R, Falk S, Faivre-Finn C, Dubec M, et al. The future of image-guided radiotherapy. Clin Oncol 2017;29:662–6.

[8] Chawla S, Kim S, Wang S, Poptani H. Diffusion-weighted imaging in head and neck cancers. Futur Oncol 2009;5:959–75.

[9] Weygand J, Fuller CD, Ibbott GS, Mohamed ASR, Ding Y, Yang J, et al. Spatial precision in magnetic resonance imaging–guided radiation therapy: the role of geometric distortion. Int J Radiat Oncol Biol Phys 2016;95:1304–16.

[10] Rong Y, Rosu-Bubulac M, Benedict SH, Cui Y, Ruo R, Connell T, et al. Rigid and Deformable Image Registration for Radiation Therapy: A Self-Study Evaluation Guide in YYYY Clinical Trial Participation. Pract Radiat Oncol 2021.

[11] Hill DLG, Batchelor PG, Holden M, Hawkes DJ. Medical image registration. Phys Med Biol 2001;46:R1.

[12] Kessler ML. Image registration and data fusion in radiation therapy. Br J Radiol 2006;79. https://doi.org/10.1259/bjr/70617164.

[13] Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: A survey. IEEE Trans Med Imaging 2013;32:1153–90.

[14] Oliveira FPM, Tavares JMRS. Medical image registration: a review. Comput Methods Biomech Biomed Engin 2014;17:73–93.

[15] Schwartz DL, Garden AS, Shah SJ, Chronowski G, Sejpal S, Rosenthal DI, et al. Adaptive radiotherapy for head and neck cancer—dosimetric results from a prospective clinical trial. Radiother Oncol 2013;106:80–4.

[16] Mohamed ASR, Ruangskul M-N, Awan MJ, Baron CA, Kalpathy-Cramer J, Castillo R, et al. Quality assurance assessment of diagnostic and radiation therapy–simulation CT image registration for head and neck radiation therapy: anatomic region of interest–based comparison of rigid and deformable algorithms. Radiology 2015;274:752–63.

[17] Kiser K, Meheissen MAM, Mohamed ASR, Kamal M, Ng SP, Elhalawani H, et al. Prospective quantitative quality assurance and deformation estimation of MRI-CT image registration in simulation of head and neck radiotherapy patients. Clin Transl Radiat Oncol 2019;18:120–7.

[18] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30:1323–41.

[19] Han X, Hibbard LS, Willcut V. An Efficient Inverse-Consistent Diffeomorphic Image Registration Method for Prostate Adaptive Radiotherapy. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2010;6367 LNCS:34–41. https://doi.org/10.1007/978-3-642-15989-3_5.

[20] Unser M, Aldroubi A, Eden M. B-spline signal processing. I. Theory. IEEE Trans Signal Process 1993;41:821–33.

[21] Nikolov S, Blackwell S, Zverovitch A, Mendes R, Livne M, De Fauw J, et al. Clinically

Applicable Segmentation of Head and Neck Anatomy for Radiotherapy: Deep Learning Algorithm Development and Validation Study. J Med Internet Res 2021;23:e26151.

[22] Wahid K, Ahmed S, He R, van Dijk L, Teuwen J, McDonald BA, et al. Development of a High-Performance Multiparametric MRI Oropharyngeal Primary Tumor Auto-Segmentation Deep Learning Model and Investigation of Input Channel Effects: Results from a Prospective Imaging Registry. MedRxiv 2021.

[23] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52:591–611.

[24] Wilcoxon F. Individual comparisons by ranking methods. Break. Stat., Springer; 1992, p. 196–202.

[25] Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 1952;47:583–621.

[26] Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009 n.d.

[27] Andre JB, Bammer R. Advanced diffusion-weighted magnetic resonance imaging techniques of the human spinal cord. Top Magn Reson Imaging TMRI 2010;21:367.

[28] Ou Y, Akbari H, Bilello M, Da X, Davatzikos C. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. IEEE Trans Med Imaging 2014;33:2039–65.

[29] Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang M-C, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 2009;46:786–802.

[30] Fabri D, Zambrano V, Bhatia A, Furtado H, Bergmann H, Stock M, et al. A quantitative

comparison of the performance of three deformable registration algorithms in radiotherapy. Z Med Phys 2013;23:279–90.

[31] De Luca M, Giannini V, Vignati A, Mazzetti S, Bracco C, Stasi M, et al. A fully automatic method to register the prostate gland on T2-weighted and EPI-DWI images. 2011 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE; 2011, p. 8029–32.

[32] Buerger C, Sénégas J, Kabus S, Carolus H, Schulz H, Agarwal H, et al. Comparing nonrigid registration techniques for motion corrected MR prostate diffusion imaging. Med Phys 2015;42:69–80.

[33] Eriksson M. Comparison of five methods for deformable, multi-modal image registration in prostate and pelvic area. Aalto Univeristy, 2015.

[34] Chen DQ, Dell'Acqua F, Rokem A, Garyfallidis E, Hayes DJ, Zhong J, et al. Diffusion weighted image co-registration: investigation of best practices. BioRxiv 2019:864108.

[35] Sherer M V, Lin D, Elguindi S, Duke S, Tan L-T, Cacicedo J, et al. Metrics to evaluate the performance of auto-segmentation for radiation treatment planning: A critical review. Radiother Oncol 2021.